練習問題解答

第 2 章

【2-1】

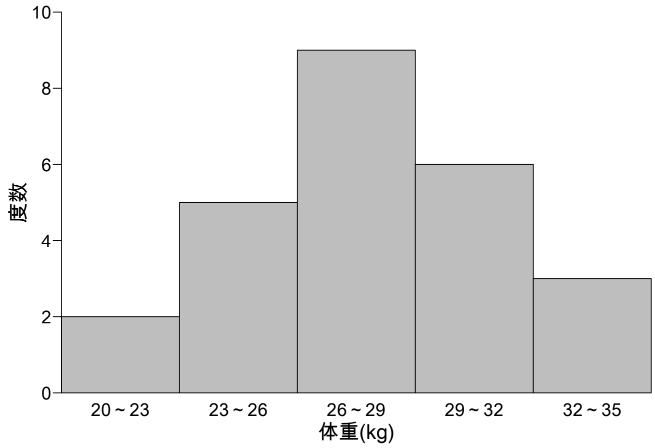
① 牛乳の脂肪含量（乳脂率）：比率尺度, 連続変数
② モース硬度：順序尺度
③ 本のページ番号（総ページ数ではない）：間隔尺度, 離散変数
④ 農家の専業・兼業の別：名義尺度
⑤ 経過時間：比率尺度, 連続変数
⑥ 時刻：間隔尺度, 連続変数
⑦ 金額：比率尺度, 離散変数
⑧ 国籍：名義尺度
⑨ 西暦年：間隔尺度, 離散変数
⑩ 成績評価（秀, 優, 良, 可, 不可）：順序尺度
⑪ 体温：間隔尺度, 連続変数
⑫ クラスの生徒数：比率尺度, 離散変数

第 3 章

【3-1】

データ数が 25 なのでスタージェスの公式より階級数の目安は 5.6 と計算され, 仮の階級数は 5 または 6 となる. データの範囲が 13.8(= 34.2 − 20.4) なので, 階級幅の目安は, 仮の階級数が 5 のときには 13.8/5 = 2.76, 仮の階級数が 6 のときには 13.8/6 = 2.30 となる. データ数が 25 とかなり小さいので大きめの値とし, 切りのよい値にまるめると, 階級幅は 3 となる. さらに, 階級分けの開始値を切りのよい値として 20 に設定すると, 終了値は 35, 階級数は 5 となる. 以上にもとづいて作成した度数分布表および度数分布図は以下のようになる.

階級（kg）（下限値〜上限値）	階級値（kg）	度数	相対度数	累積度数	累積相対度数
20 〜 23	21.5	2	0.08	2	0.08
23 〜 26	24.5	5	0.20	7	0.28
26 〜 29	27.5	9	0.36	16	0.64
29 〜 32	30.5	6	0.24	22	0.88
32 〜 35	33.5	3	0.12	25	1.00
合計	—	25	1.00	—	—



度数分布からデータ分布の概要として次のことが読み取れる. ①子牛の生時体重は階級値で 21.5 〜 33.5 kg の範囲にある. ② 26 〜 29 kg の体重で生まれる子牛が最も多く, この階級を中心として体重分布はほぼ左右対称である. ③ 80% の子牛が 23 〜 32 kg の体重で生まれる.

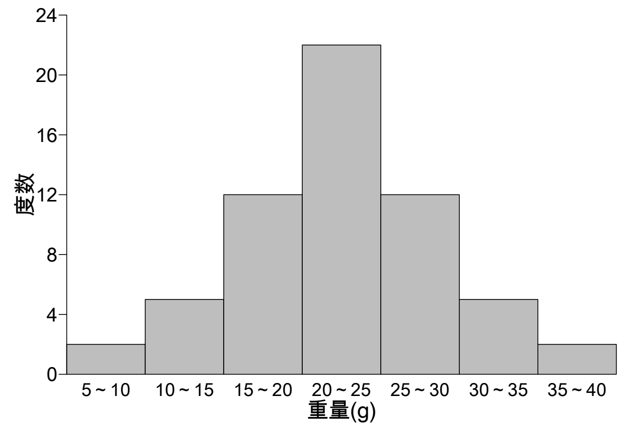
生データおよび度数分布表を利用して求められる統計量は以下のようである。

統計量	生データ	度数分布表
最小値（kg）	20.4	20
最大値（kg）	34.2	35
範囲（kg）	13.8	15
平均値（kg）	27.8	27.9
中央値（kg）	27.8	27.5
最頻値（kg）	—	27.5
平方和（kg^2）	289.1	275.8
分散（kg^2）	11.57	11.03
標準偏差（kg）	3.40	3.32
変動係数	0.123	0.119

【3-2】

データ数が 60 なので, スタージェスの公式より階級数の目安は 6.9 と計算され, 仮の階級数は 6 または 7 となる. データの範囲が $32 (= 38 - 6)$ なので, 階級幅の目安は, 仮の階級数が 6 のときには $32/6 = 5.3$, 仮の階級数が 7 のときには $32/7 = 4.6$ となり, 切りのよい値にまるめると, 階級幅は 5 となる. さらに, 階級分けの開始値を切りのよい値として 5 に設定すると, 終了値は 40, 階級数は 7 となる. 以上にもとづいて作成した度数分布表および度数分布図は以下のようになる.

階級（g）（下限値〜上限値）	階級値（g）	度数	相対度数	累積度数	累積相対度数
5 〜 10	7.5	2	0.03	2	0.03
10 〜 15	12.5	5	0.08	7	0.12
15 〜 20	17.5	12	0.20	19	0.32
20 〜 25	22.5	22	0.37	41	0.68
25 〜 30	27.5	12	0.20	53	0.88
30 〜 35	32.5	5	0.08	58	0.97
35 〜 40	37.5	2	0.03	60	1.00
合計	—	60	1.00	—	—



度数分布からデータ分布の概要として次のことが読み取れる. ①ミニトマトの重量は階級値で 7.5 〜 37.5 g の範囲にある. ② 20 〜 25 g の果実が最も多く, この階級を中心として重量分布は左右対称である. ③ 77% の果実が 15 〜 30 g である.

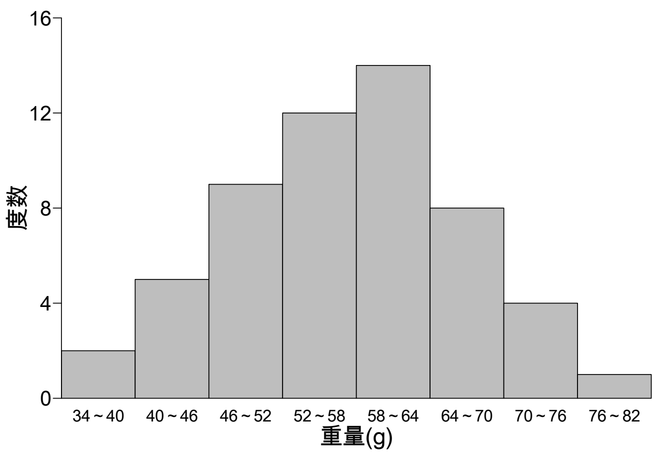
生データおよび度数分布表を利用して求められる統計量は以下のようである.

統計量	生データ	度数分布表
最小値（g）	6	5
最大値（g）	38	40
範囲（g）	32	35
平均値（g）	22.3	22.5
中央値（g）	22	22.5
最頻値（g）	22	22.5
平方和（g^2）	2762.2	2500.0
分散（g^2）	46.04	41.67
標準偏差（g）	6.79	6.45
変動係数	0.304	0.287

【3-3】

度数分布表および度数分布図は以下のようになる．

階級（g）（下限値〜上限値）	階級値（g）	度数	相対度数	累積度数	累積相対度数
34 〜 40	37	2	0.04	2	0.04
40 〜 46	43	5	0.09	7	0.13
46 〜 52	49	9	0.16	16	0.29
52 〜 58	55	12	0.22	28	0.51
58 〜 64	61	14	0.25	42	0.76
64 〜 70	67	8	0.15	50	0.91
70 〜 76	73	4	0.07	54	0.98
76 〜 82	79	1	0.02	55	1.00
合計	—	55	1.00	—	—



度数分布からデータ分布の概要として次のことが読み取れる．①鶏卵重量は階級値で 37 〜 79 g の範囲にある．② 58 〜 64 g（M 等級）の卵が最も多く，この階級を中心として重量分布は左右非対称で，データが重量の大きい方よりも小さい方へ広がっており，重量が大きい卵よりも小さい卵の方が多い．③ 47% の卵が 52 〜 64 g（MS もしくは M 等級），78% の卵が 46 〜 70 g（S 〜 L 等級）である．

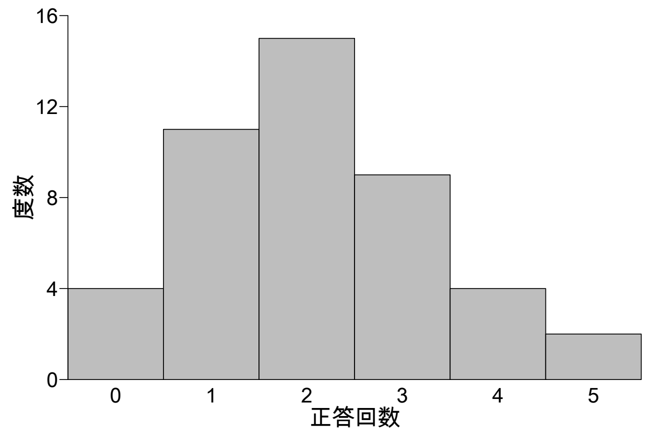
生データおよび度数分布表を利用して求められる統計量は以下のようである．

統計量	生データ	度数分布表
最小値（g）	38	34
最大値（g）	80	82
範囲（g）	42	48
平均値（g）	56.8	57.3
中央値（g）	57	55
最頻値（g）	60	61
平方和（g^2）	4933.5	4931.3
分散（g^2）	91.36	91.32
標準偏差（g）	9.56	9.56
変動係数	0.168	0.167

【3-4】

度数分布表および度数分布図は以下のようになる．

階級（回）	階級値（回）	度数	相対度数	累積度数	累積相対度数
0	0	4	0.09	4	0.09
1	1	11	0.24	15	0.33
2	2	15	0.33	30	0.67
3	3	9	0.20	39	0.87
4	4	4	0.09	43	0.96
5	5	2	0.04	45	1.00
合計	―	45	1.00	―	―



度数分布からデータ分布の概要として次のことが読み取れる．①正答回数は 0 〜 5 の範囲にある．②正答回数 2 の被験者が最も多く，この階級を中心として正答回数の分布は左右非対称で，データが正答回数の小さい方よりも大きい方へ広がっている．③ 78% の被験者が 1, 2 もしくは 3 回の正答回数である．

生データおよび度数分布表を利用して求められる統計量は以下のようであり，データがとり得る個々の値を階級（階級値）としているため，生データと度数分布表のいずれを利用しても統計量の値は等しい．

統計量	生データ	度数分布表
最小値（回）	0	0
最大値（回）	5	5
範囲（回）	5	5
平均値（回）	2.1	2.1
中央値（回）	2	2
最頻値（回）	2	2
平方和（回^2）	69.6	69.6
分散（回^2）	1.55	1.55
標準偏差（回）	1.24	1.24
変動係数	0.596	0.596

第 4 章

【4-1】

① $2 \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{18}$　② $3 \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{12}$

③ $5 \times \frac{1}{6} \times \frac{1}{6} = \frac{5}{36}$　④ $6 \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{6}$

⑤ $1 - 6 \times \frac{1}{6} \times \frac{1}{6} = \frac{5}{6}$

【4-2】

期待値 $E(x)$ は $100 \times \frac{1}{80} + 50 \times \frac{9}{80} + 30 \times \frac{7}{80} = 1.25 + 5.625 + 2.625 = 9.5$ 点となる.

【4-3】

2 項分布 $P(X = x) = {}_n\mathrm{C}_x\, p^x (1-p)^{n-x}$ に $n = 4,\ x = 0, 1, 2, 3, 4,\ p = 1/6$ を適用する.

	X					合計
	0	1	2	3	4	
確率	0.4823	0.3858	0.1157	0.0154	0.0008	1

【4-4】

1 人の患者に新薬を試みて効果があれば「成功」, 効果がなければ「失敗」とする. 個々の患者への効果は独立であり, かつ, 同じ効果を期待できるとすれば, 「$X =$ 効果が認められた患者数」は $n = 20, p = 0.3$ の 2 項分布に従う.

(1) $P(X = 8) = {}_{20}\mathrm{C}_8 \times 0.3^8 \times (1 - 0.3)^{12} = 0.1144$

(2) $P(8 \leqq X \leqq 12) = \sum_{i=8}^{12} P(X = i) = 0.1144 + 0.0654 + 0.0308 + 0.0120 + 0.0039 = 0.2265$

(3) $P(X \geqq 10) = 1 - P(X \leqq 9) = 1 - \sum_{i=0}^{9} P(X = i) = 1 - 0.9520 = 0.0480$

【4-5】

(1) $P(x = 0) = \frac{4^0}{0!} e^{-4} = e^{-4} = 0.0183$

(2) $P(x = 1) = \frac{4^1}{1!} e^{-4} = 4e^{-4} = 4 \times 0.01832 = 0.0733$

(3) $P(x \geqq 2) = 1 - \{P(x = 0) + P(x = 1)\} = 1 - 0.0183 - 0.0733 = 0.9084$

【4-6】

(1) $P(Z \geqq 1.04) = 0.14917$

(2) $P(0 \leqq Z \leqq 0.73) = 0.5 - P(Z \geqq 0.73) = 0.5 - 0.23270 = 0.26730$

(3) $P(0.41 \leqq Z \leqq 1.23) = \{0.5 - P(Z \geqq 1.23)\} - \{0.5 - P(Z \geqq 0.41)\}$
$= (0.5 - 0.10935) - (0.5 - 0.34090) = 0.39065 - 0.15910 = 0.23155$

(4) $P(-1.11 \leqq Z \leqq 0.77) = \{0.5 - P(Z \geqq 1.11)\} + \{0.5 - P(Z \geqq 0.77)\}$
$= (0.5 - 0.13350) + (0.5 - 0.22065) = 0.3665 + 0.27935 = 0.64585$

【4-7】

(1) $P(X \geqq 4.7) = P\left(\dfrac{X-3}{2.5} \geqq \dfrac{4.7-3}{2.5}\right) = P(Z \geqq 0.68) = 0.24825$

(2) $P(3 \leqq X \leqq 4.15) = P\left(\dfrac{3-3}{2.5} \leqq \dfrac{X-3}{2.5} \leqq \dfrac{4.15-3}{2.5}\right) = P(0 \leqq Z \leqq 0.46)$
$= 0.5 - P(Z \geqq 0.46) = 0.5 - 0.32276 = 0.17724$

(3) $P(-3.3 \leqq X \leqq 0.2) = P\left(\dfrac{-3.3-3}{2.5} \leqq \dfrac{X-3}{2.5} \leqq \dfrac{0.2-3}{2.5}\right) = P(-2.52 \leqq Z \leqq -1.12)$
$= \{0.5 - P(Z \geqq 2.52)\} - \{0.5 - P(Z \geqq 1.12)\} = (0.5 - 0.00587) - (0.5 - 0.13136)$
$= 0.49413 - 0.36864 = 0.12549$

(4) $P(-0.95 \leqq X \leqq 7.55) = \left(\dfrac{-0.95-3}{2.5} \leqq \dfrac{X-3}{2.5} \leqq \dfrac{7.55-3}{2.5}\right) = P(-1.58 \leqq Z \leqq 1.82)$
$= \{0.5 - P(Z \geqq 1.58)\} + \{0.5 - P(Z \geqq 1.82)\} = (0.5 - 0.05705) + (0.5 - 0.03438)$
$= 0.44295 + 0.46562 = 0.90857$

【4-8】

(1)70 g 以上の果実の割合

品種 A の場合, 重量 X は $N(50, 100)$ に従うので, $Z = (X - 50)/10$ は $N(0, 1)$ に従う. X が 70 以上となる確率 $P(X \geqq 70)$ は, X と 70 を標準化すると, Z が 2 以上となる確率 $P(Z \geqq 2)$ となり, 標準正規分布表（付表 2）より 0.02275 が得られる.

$$P(X \geqq 70) = P\left(\frac{X-50}{10} \geqq \frac{70-50}{10}\right) = P(Z \geqq 2) = 0.02275$$

ゆえに, 70 g 以上の果実は全体の 2.275%（約 2.3%）となる.

品種 B の場合, 重量 X は $N(50, 225)$ に従うので, $Z = (X - 50)/15$ は $N(0, 1)$ に従う. X が 70 以上となる確率 $P(X \geqq 70)$ は, X と 70 を標準化すると, Z が 1.33 以上となる確率 $P(Z \geqq 1.33)$ となり, 標準正規分布表（付表 2）より 0.09176 が得られる.

$$P(X \geqq 70) = P\left(\frac{X-50}{15} \geqq \frac{70-50}{15}\right) = P(Z \geqq 1.33) = 0.09176$$

ゆえに, 70 g 以上の果実は全体の 9.176%（約 9.2%）となる.

(2) 上位 10% に入る果実の重量

いずれの品種も $P(X \geqq a) = 0.1$ となる a を求めればよい. 品種 A の場合, X と a を標準化すると,

$$P(X \geqq a) = P\left(\frac{X-50}{10} \geqq \frac{a-50}{10}\right) = P\left(Z \geqq \frac{a-50}{10}\right) = 0.1$$

となる. 標準正規分布表（付表 2）より上側確率 0.1 を与える z 値は 1.28（0.1 に最も近い 0.10027 に対応する z 値を採用する）なので,

$$\frac{a-50}{10} = 1.28$$

となり, $a = 1.28 \times 10 + 50 = 62.8$ が得られる.

ゆえに, 上位 10% に入る果実は 62.8 g 以上である.

品種 B の場合, X と a を標準化すると,

$$P(X \geqq a) = P\left(\frac{X-50}{15} \geqq \frac{a-50}{15}\right) = P\left(Z \geqq \frac{a-50}{15}\right) = 0.1$$

となる。品種 A の場合と同じく，上側確率 0.1 を与える z 値は 1.28 なので,

$$\frac{a-50}{15}=1.28$$

となり, $a=1.28\times 15+50=69.2$ が得られる．

ゆえに, 上位 10% に入る果実は 69.2 g 以上である.

厳密には, 0.1 に最も近い確率として 0.10027 を用いているので, 計算される重量は, 両品種とも, 上位 10.027% に入る重量である.

【4-9】

重量 X は $N(2000,40^2)$ に従うので, $Z=(X-2000)/40$ は $N(0,1)$ に従う．X が 1950 以下となる確率 $P(X\leqq 1950)$ は, X と 1950 を標準化すると, Z が -1.25 以下となる確率 $P(Z\leqq -1.25)$ となり, 標準正規分布表（付表 2）より 0.10565 が得られる.

$$P(X\leqq 1950)=P\left(\frac{X-2000}{40}\leqq\frac{1950-2000}{40}\right)=P(Z\leqq -1.25)=P(Z\geqq 1.25)=0.10565$$

ゆえに, 1950 g 以下の袋は 100 袋中約 11 袋となる（袋の数は整数なので 10.565 を整数にまるめる）.

整数へのまるめに四捨五入を用いると, この解答のように, 数値が切り上げられる場合には, 得られる値は条件を確実に満たす数よりも 1 大きくなる．すなわち, 1950 g 以下の袋は 100 袋中 10.565 袋なので, 四捨五入により得られる 11 袋のうち, 条件を確実に満たすのは 10 袋である．残りの 1 袋は, 条件を確実には満たさないが, 満たす確率は 0.5 以上である．四捨五入によるまるめは, 条件を満たす最も近い数を与えることになる.

第 5 章

【5-1】

データ $x_i (i = 1, 2, \cdots, n; n = 20)$ の総和および 2 乗和を計算し, 平方和 S, 不偏分散 s^2, 標準偏差 s および自由度 f を求めておく.

$$\sum_{i=1}^{n} x_i = 81 + 95 + \cdots + 92 = 1860, \qquad \sum_{i=1}^{n} x_i^2 = 81^2 + 95^2 + \cdots + 92^2 = 177176$$
$$S = 177176 - \frac{1}{20} \times 1860^2 = 4196, \qquad s^2 = 4196/19 = 220.84$$
$$s = \sqrt{220.84} = 14.86, \qquad f = 20 - 1 = 19$$

(1) 体重の母平均の推定

①点推定

$$\mu\text{の推定値} = \bar{x} = 1860/20 = 93.0$$

となり, 養成池で飼育されている魚の平均体重は 93.0 g と推定される.

②区間推定

信頼係数 95% を採用すると, 付表 3 より $t(19, 0.025) = 2.093$ なので,

$$\text{下側 95\% 信頼限界} = 93.0 - 2.093 \times 14.86/\sqrt{20} = 86.05$$
$$\text{上側 95\% 信頼限界} = 93.0 + 2.093 \times 14.86/\sqrt{20} = 99.95$$
$$\text{95\% 信頼区間} : 86.05 \leqq \mu \leqq 99.95$$

となり, 信頼係数 95% で, 養成池で飼育されている魚の平均体重は 86.05 〜 99.95 g の範囲にあるといえる.

また, 信頼係数 99% を採用すると, 付表 3 より $t(19, 0.005) = 2.861$ なので,

$$\text{下側 99\% 信頼限界} = 93.0 - 2.861 \times 14.86/\sqrt{20} = 83.49$$
$$\text{上側 99\% 信頼限界} = 93.0 + 2.861 \times 14.86/\sqrt{20} = 102.51$$
$$\text{99\% 信頼区間} : 83.49 \leqq \mu \leqq 102.51$$

となり, 信頼係数 99% で, 養成池で飼育されている魚の平均体重は 83.49 〜 102.51 g の範囲にあるといえる.

(2) 体重の母平均の検定（$\mu_0 = 100$）

①仮説の設定

$$\mathrm{H}_0 : \mu = 100, \qquad \mathrm{H}_1 : \mu \neq 100$$

②統計量 t の計算

$$t = \frac{93.0 - 100.0}{14.86/\sqrt{20}} = -2.107$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると, 付表 3 より $t(19, 0.025) = 2.093$ が得られ, 棄却域は $t < -2.093$ および $t > 2.093$ となる. t の値 $(= -2.107)$ が下側有意点（-2.093）よりも小さく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する. すなわち, 有意水準 0.05 で, 養成池で飼育されている魚の平均体重は 100 g と異なるといえる.

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.01 を採用すると, 付表 3 より $t(19, 0.005) = 2.861$ が得られ, 棄却域は $t < -2.861$ お

よび $t > 2.861$ となる．t の値（$= -2.107$）が下側有意点（-2.861）と上側有意点（2.861）の間に位置し, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 養成池で飼育されている魚の平均体重は 100 g と異ならないといえる．

(3) 体重の母分散の推定

①点推定

$$\sigma^2 \text{の推定値} = s^2 = 220.84$$

となり, 養成池で飼育されている魚の体重の分散は $220.84\,\mathrm{g}^2$ と推定される．

②区間推定

信頼係数 95% を採用すると, 付表 4 より $\chi^2(19, 0.025) = 32.852$, $\chi^2(19, 0.975) = 8.907$ なので,

$$\text{下側 95\% 信頼限界} = 4196/32.852 = 127.72$$
$$\text{上側 95\% 信頼限界} = 4196/8.907 = 471.09$$
$$\text{95\% 信頼区間} : 127.72 \leqq \sigma^2 \leqq 471.09$$

となり, 信頼係数 95% で, 養成池で飼育されている魚の体重の分散は $127.72 \sim 471.09\,\mathrm{g}^2$ の範囲にあるといえる．

また, 信頼係数 99% を採用すると, 付表 4 より $\chi^2(19, 0.005) = 38.582$, $\chi^2(19, 0.995) = 6.844$ なので,

$$\text{下側 99\% 信頼限界} = 4196/38.582 = 108.76$$
$$\text{上側 99\% 信頼限界} = 4196/6.844 = 613.09$$
$$\text{99\% 信頼区間} : 108.76 \leqq \sigma^2 \leqq 613.09$$

となり, 信頼係数 99% で, 養成池で飼育されている魚の体重の分散は $108.76 \sim 613.09\,\mathrm{g}^2$ の範囲にあるといえる．

(4) 体重の母分散の検定（$\sigma_0^2 = 121$）

①仮説の設定

$$\mathrm{H}_0 : \sigma^2 = 121, \qquad \mathrm{H}_1 : \sigma^2 \neq 121$$

②統計量 χ^2 の計算

$$\chi^2 = 4196/121 = 34.68$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると, 付表 4 より $\chi^2(19, 0.975) = 8.907$, $\chi^2(19, 0.025) = 32.852$ が得られ, 棄却域は $\chi^2 < 8.907$ および $\chi^2 > 32.852$ となる．χ^2 の値（$= 34.68$）が上側有意点（32.852）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する．すなわち, 有意水準 0.05 で, 養成池で飼育されている魚の体重の分散は $121\,\mathrm{g}^2$ と異なるといえる．

④ 有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.01 を採用すると, 付表 4 より $\chi^2(19, 0.995) = 6.844$, $\chi^2(19, 0.005) = 38.582$ が得られ, 棄却域は $\chi^2 < 6.844$ および $\chi^2 > 38.582$ となる．χ^2 の値（$= 34.68$）が下側有意点（6.844）と上側有意点（38.582）の間にあり, 棄却域には含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 養成池で飼育されている魚の体重の分散は $121\,\mathrm{g}^2$ と異ならないといえる．

【5-2】

品種 A のデータを x, 品種 B のデータを y とし, d_i（$d_i = x_i - y_i\,; i = 1, 2, \cdots, n; n = 10$）の総和および 2 乗和を計算し, 標準偏差 s_d および自由度 f を求めておく. また, 念のために $\bar{x}$ と $\bar{y}$ も計算する.

$$\sum_{i=1}^{n} d_i = 53 - 7 + \cdots + 125 = 492, \qquad \sum_{i=1}^{n} d_i^2 = 53^2 + (-7)^2 + \cdots + 125^2 = 46754$$

$$s_d = \sqrt{\frac{1}{9} \times \left(46754 - \frac{1}{10} \times 492^2\right)} = \sqrt{2505.29} = 50.05, \qquad f = 10 - 1 = 9$$

$$\bar{x} = (884 + 840 + \cdots + 1250)/10 = 7950/10 = 795.0$$

$$\bar{y} = (831 + 847 + \cdots + 1125)/10 = 7458/10 = 745.8$$

(1) 収量の差の母平均の推定

①点推定

$$\mu_d \text{の推定値} = \bar{d} = 492/10 = 49.2$$

となり, 品種 A と B の収量の差の母平均は 49.2 g/m^2 と推定される. このことは $\bar{x} - \bar{y} = 795.0 - 745.8 = 49.2$ と一致することで確認できる.

②区間推定

信頼係数 95% を採用すると, 付表 3 より $t(9, 0.025) = 2.262$ なので,

$$\text{下側 95\% 信頼限界} = 49.2 - 2.262 \times 50.05/\sqrt{10} = 13.40$$

$$\text{上側 95\% 信頼限界} = 49.2 + 2.262 \times 50.05/\sqrt{10} = 85.00$$

$$\text{95\% 信頼区間} : 13.40 \leqq \mu_d \leqq 85.00$$

となり, 信頼係数 95% で, 品種 A と B の収量の差の母平均は 13.40 〜 85.00 g/m^2 の範囲にあるといえる.

また, 信頼係数 99% を採用すると, 付表 3 より $t(9, 0.005) = 3.250$ なので,

$$\text{下側 99\% 信頼限界} = 49.2 - 3.250 \times 50.05/\sqrt{10} = -2.24$$

$$\text{上側 99\% 信頼限界} = 49.2 + 3.250 \times 50.05/\sqrt{10} = 100.64$$

$$\text{99\% 信頼区間} : -2.24 \leqq \mu_d \leqq 100.64$$

となり, 信頼係数 99% で, 品種 A と B の収量の差の母平均は -2.24 〜 100.64 g/m^2 の範囲にあるといえる.

(2) 収量の差の母平均の検定（$\mu_0 = 0$ の両側検定）

①仮説の設定

$$\mathrm{H}_0 : \mu_d = 0, \qquad \mathrm{H}_1 : \mu_d \neq 0$$

②統計量 t の計算

$$t = \frac{49.2 - 0}{50.05/\sqrt{10}} = 3.108$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると, 付表 3 より $t(9, 0.025) = 2.262$ が得られ, 棄却域は $t < -2.262$ および $t > 2.262$ となる. t の値 $(= 3.108)$ が上側有意点（2.262）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する. すなわち, 有意水準 0.05 で, 品種 A と B の収量の差の母平均は 0 と異なる（品種 A と B によって収量が異なる）といえる.

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.01 を採用すると, 付表 3 より $t(9, 0.005) = 3.250$ が得られ, 棄却域は $t < -3.250$ および $t > 3.250$ となる．t の値（$= 3.108$）が下側有意点（-3.250）と上側有意点（3.250）の間に位置し, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 品種 A と B の収量の差の母平均は 0 と異ならない（品種 A と B によって収量は異ならない）といえる．

(3) 収量の差の母平均の検定（$\mu_0 = 0$ の片側検定）

①仮説の設定

2 つの品種の収量の差 d_i が品種 A の収量から品種 B の収量を引いたものであり, 新品種 A の方が従来品種 B よりも収量が高いと予想されるので,

$$\mathrm{H}_0 : \mu_d = 0, \qquad \mathrm{H}_1 : \mu_d > 0$$

②統計量 t の計算

(2) と同じ（$t = 3.108$）．

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.05 を採用すると, 付表 3 より $t(9, 0.05) = 1.833$ が得られ, 棄却域は $t > 1.833$ となる．t の値 ($= 3.108$) が有意点（1.833）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する．すなわち, 有意水準 0.05 で, 品種 A と B の収量の差の母平均は 0 よりも大きい（品種 A の方が品種 B よりも収量が高い）といえる．

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.01 を採用すると, 付表 3 より $t(9, 0.01) = 2.821$ が得られ, 棄却域は $t > 2.821$ となる．t の値（$= 3.108$）が有意点（2.821）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する．すなわち, 有意水準 0.01 でも, 品種 A と B の収量の差の母平均は 0 よりも大きい（品種 A の方が品種 B よりも収量が高い）といえる．

(4) 収量の差の母平均の検定（$\mu_0 = 20$ の片側検定）

①仮説の設定

$$\mathrm{H}_0 : \mu_d = 20, \qquad \mathrm{H}_1 : \mu_d > 20$$

②統計量 t の計算

$$t = \frac{49.2 - 20}{50.05/\sqrt{10}} = 1.845$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.05 を採用すると, 付表 3 より $t(9, 0.05) = 1.833$ が得られ, 棄却域は $t > 1.833$ となる．t の値（$= 1.845$）が有意点（1.833）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する．すなわち, 有意水準 0.05 で, 品種 A と B の収量の差の母平均は $20\,\mathrm{g/m^2}$ を超える（品種 A は品種 B よりも $20\,\mathrm{g/m^2}$ を超える高収量である）といえる．

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.01 を採用すると, 付表 3 より $t(9, 0.01) = 2.821$ が得られ, 棄却域は $t > 2.821$ となる．t の値（$= 1.845$）が有意点（2.821）よりも小さく, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 品種 A と B の収量の差の母平均は $20\,\mathrm{g/m^2}$ と異ならない（品種 A は品種 B よりも $20\,\mathrm{g/m^2}$ だけ高収量である）といえる．

【5-3】

品種 A のデータを x, 品種 B のデータを y とし, データ x_i（$i = 1, 2, \cdots, n_x; n_x = 8$）および y_i（$i = 1, 2, \cdots, n_y; n_y = 7$）の総和および 2 乗和を計算し, 平方和 S_x, S_y, 不偏分散 s_x^2, s_y^2 および自由度 f_1, f_2 を求めておく.

$$\sum_{i=1}^{n_x} x_i = 1120 + 1050 + \cdots + 1060 = 8480, \qquad \sum_{i=1}^{n_y} y_i = 900 + 1030 + \cdots + 840 = 6720$$
$$\sum_{i=1}^{n_x} x_i^2 = 1120^2 + 1050^2 + \cdots + 1060^2 = 9034000$$
$$\sum_{i=1}^{n_y} y_i^2 = 900^2 + 1030^2 + \cdots + 840^2 = 6488000$$
$$S_x = 9034000 - \frac{1}{8} \times 8480^2 = 45200, \qquad S_y = 6488000 - \frac{1}{7} \times 6720^2 = 36800$$
$$s_x^2 = 45200/7 = 6457.14, \qquad s_y^2 = 36800/6 = 6133.33$$
$$f_1 = 8 - 1 = 7, \qquad f_2 = 7 - 1 = 6$$

(1) 収量の母分散の違い（分散比）の検定

①仮説の設定

$$\mathrm{H}_0 : \sigma_x/\sigma_y = 1, \qquad \mathrm{H}_1 : \sigma_x/\sigma_y \neq 1$$

②統計量 F の計算

$s_x^2 > s_y^2$ なので, 品種 A, B と x, y との対応は以上のままでよく, 統計量 F を s_x^2/s_y^2 として計算する.

$$F = 6457.14/6133.33 = 1.053$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると , 付表 5 より $F(7, 6, 0.025) = 5.695$ が得られ, 上側棄却域は $F > 5.695$ となる. F の値（$= 1.053$）は上側有意点（5.695）よりも小さく, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する. すなわち, 有意水準 0.05 で, 品種 A と B の収量の母分散は異ならないといえる.

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

有意水準 0.05 で H_0 が採択されるので, 有意水準 0.01 でも H_0 は採択され, 品種 A と B の収量の母分散は異ならないといえる. ちなみに, 上側棄却域は $F > F(7, 6, 0.005) = 10.786$ である.

品種 A と B の収量の母分散が異ならないとみなされるので, 以下では, 2 つの母集団（品種 A と B）の母分散が等しい場合の推定と検定を用いる. 2 つの標本平均 $\bar{x}, \bar{y}$, 共通の標準偏差 s_c および自由度（平均値の差の自由度）f を求めておく.

$$\bar{x} = 8480/8 = 1060, \qquad \bar{y} = 6720/7 = 960$$
$$s_c = \sqrt{\frac{45200 + 36800}{(8-1) + (7-1)}} = \sqrt{6307.69} = 79.421$$
$$f = (8-1) + (7-1) = 8 + 7 - 2 = 13$$

(2) 収量の母平均の差の推定

①点推定

$$\mu_x - \mu_y \text{の推定値} = \bar{x} - \bar{y} = 1060 - 960 = 100$$

となり, 品種 A と B の収量の母平均の差は $100\,\mathrm{g/m^2}$ と推定される.

②区間推定

信頼係数 95% を採用すると, 付表 3 より $t(13, 0.025) = 2.160$ なので,

$$\text{下側 95\% 信頼限界} = 100 - 2.160 \times 79.421 \times \sqrt{1/8 + 1/7} = 11.21$$
$$\text{上側 95\% 信頼限界} = 100 + 2.160 \times 79.421 \times \sqrt{1/8 + 1/7} = 188.79$$
$$\text{95\% 信頼区間} : 11.21 \leqq \mu_x - \mu_y \leqq 188.79$$

となり, 信頼係数 95% で, 品種 A と B の収量の母平均の差は $11.21 \sim 188.79\,\mathrm{g/m^2}$ の範囲にあるといえる.

また, 信頼係数 99% を採用すると, 付表 3 より $t(13, 0.005) = 3.012$ なので,

$$\text{下側 99\% 信頼限界} = 100 - 3.012 \times 79.421 \times \sqrt{1/8 + 1/7} = -23.81$$
$$\text{上側 99\% 信頼限界} = 100 + 3.012 \times 79.421 \times \sqrt{1/8 + 1/7} = 223.81$$
$$\text{99\% 信頼区間} : -23.81 \leqq \mu_x - \mu_y \leqq 223.81$$

となり, 信頼係数 99% で, 品種 A と B の収量の母平均の差は $-23.81 \sim 223.81\,\mathrm{g/m^2}$ の範囲にあるといえる.

(3) 収量の母平均の差の検定（$\mu_0 = 0$ の両側検定）

①仮説の設定

$$\mathrm{H}_0 : \mu_x - \mu_y = 0, \qquad \mathrm{H}_1 : \mu_x - \mu_y \neq 0$$

②統計量 t の計算

$$t = \frac{100 - 0}{79.421 \times \sqrt{1/8 + 1/7}} = 2.433$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると, 付表 3 より $t(13, 0.025) = 2.160$ が得られ, 棄却域は $t < -2.160$ および $t > 2.160$ となる. t の値（$= 2.433$）が上側有意点（2.160）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する. すなわち, 有意水準 0.05 で, 品種 A と B の収量の母平均の差は 0 と異なる（品種 A と B によって収量が異なる）といえる.

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.01 を採用すると, 付表 3 より $t(13, 0.005) = 3.012$ が得られ, 棄却域は $t < -3.012$ および $t > 3.012$ となる. t の値（$= 2.433$）が下側有意点（-3.012）と上側有意点（3.012）の間に位置し, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する. すなわち, 有意水準 0.01 では, 品種 A と B の収量の母平均の差は 0 と異ならない（品種 A と B によって収量は異ならない）といえる.

(4) 収量の母平均の差の検定（$\mu_0 = 0$ の片側検定）

①仮説の設定　新品種 A の方が従来品種 B よりも収量が高いと予想されるので,

$$\mathrm{H}_0 : \mu_x - \mu_y = 0, \qquad \mathrm{H}_1 : \mu_x - \mu_y > 0$$

②統計量 t の計算

(3) と同じ（$t = 2.433$）.

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.05 を採用すると, 付表 3 より $t(13, 0.05) = 1.771$ が得られ, 棄却域は $t > 1.771$ となる. t の値（$= 2.433$）が有意点（1.771）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択す

る．すなわち, 有意水準 0.05 で, 品種 A と B の収量の母平均の差は 0 よりも大きい（品種 A の方が品種 B よりも収量が高い）といえる．

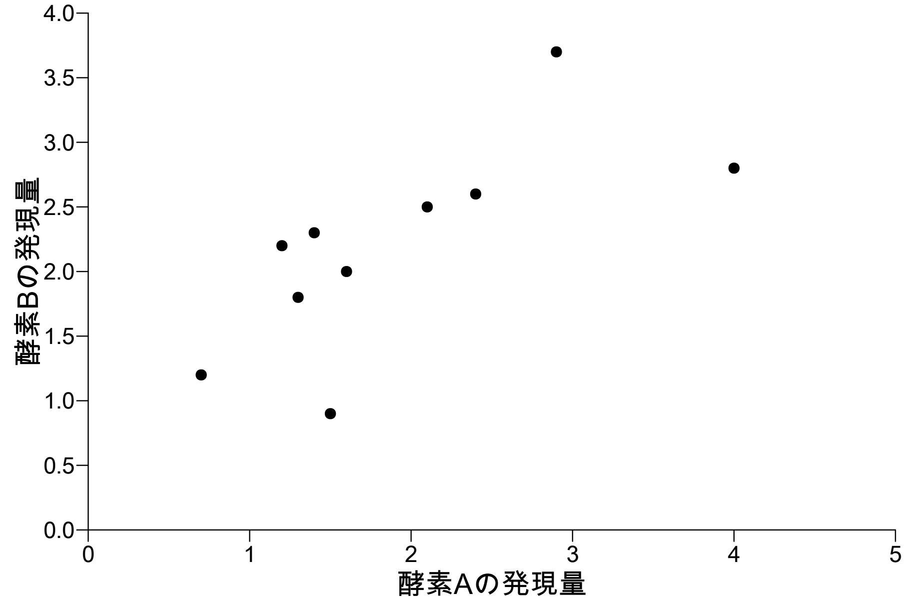
④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

上側の片側検定で有意水準 0.01 を採用すると, 付表 3 より $t(13, 0.01) = 2.650$ が得られ, 棄却域は $t > 2.650$ となる．t の値（$= 2.433$）が有意点（2.650）よりも小さく, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 品種 A と B の収量の母平均の差は 0 と異ならない（品種 A と B によって収量は異ならない）といえる．

第 6 章

【6-1】

(1) 散布図の作成



酵素 A と酵素 B の発現量の関係

(2) 相関係数の計算

$$n = 10$$

$$s_x = \sqrt{\frac{1}{10-1}\left(44.97 - \frac{1}{10} \times 19.1^2\right)} = 0.9712$$

$$s_y = \sqrt{\frac{1}{10-1}\left(54.16 - \frac{1}{10} \times 22.0^2\right)} = 0.8$$

$$s_{xy}^2 = \frac{1}{10-1}\left(47.01 - \frac{1}{10} \times 19.1 \times 22.0\right) = 0.5544$$

$$r = \frac{0.5544}{0.9712 \times 0.8} = 0.7136$$

(3) 相関係数の検定

①仮説の設定

$$\mathrm{H}_0 : \rho = 0, \qquad \mathrm{H}_1 : \rho \neq 0$$

②統計量 t の計算

$$t = \frac{0.7136 \times \sqrt{10-2}}{\sqrt{1 - 0.7136^2}} = 2.881, \qquad f = 10 - 2 = 8$$

③有意水準 0.05 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.05 を採用すると, 付表 3 より $t(8, 0.025) = 2.306$ が得られ, 棄却域は $t < -2.306$ および $t > 2.306$ となる. t の値（$= 2.881$）が上側有意点（2.306）よりも大きく, 棄却域に含まれるので, H_0 を棄却し, H_1 を採択する. すなわち, 有意水準 0.05 で, 酵素 A の発現量と酵素 B の発現量の間に相関関係があ

るといえる．この結論は, 相関係数 $r = 0.7136$ が, 付表 6 における $f = 8$, 両側確率 $p = 0.05$ に対応する値 0.6319 よりも大きいことからも導かれる．

④有意水準 0.01 での棄却域の設定, 統計量との照合, 仮説の検証および結論

両側検定で有意水準 0.01 を採用すると, 付表 3 より $t(8, 0.005) = 3.355$ が得られ, 棄却域は $t < -3.355$ および $t > 3.355$ となる．t の値（$= 2.881$）が下側有意点（-3.355）と上側有意点（3.355）の間に位置し, 棄却域に含まれないので, H_0 を採択し, H_1 を棄却する．すなわち, 有意水準 0.01 では, 酵素 A の発現量と酵素 B の発現量の間に相関関係はないといえる．この結論は, 相関係数 $r = 0.7136$ が, 付表 6 における $f = 8$, 両側確率 $p = 0.01$ に対応する値 0.7646 よりも小さいことからも導かれる．

【6-2】

(1) 回帰式 $y = a + bx$ の計算（a と b の推定）

$$n = 10, \qquad \bar{x} = 2462/10 = 246.2, \qquad \bar{y} = 973/10 = 97.3$$

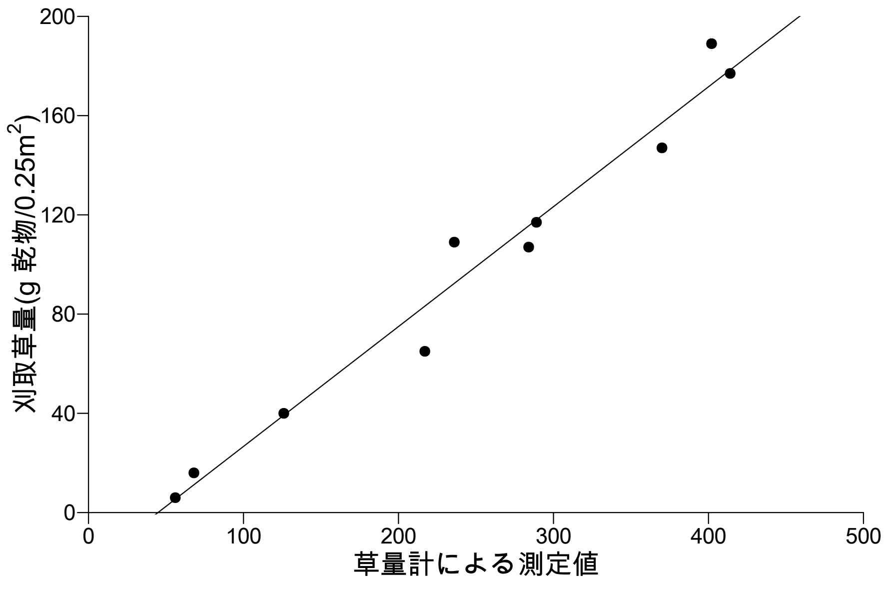
$$s_x^2 = \frac{1}{10-1}\left(760498 - \frac{1}{10} \times 2462^2\right) = 17150.4$$

$$s_{xy}^2 = \frac{1}{10-1}\left(314140 - \frac{1}{10} \times 2462 \times 973\right) = 8287.5$$

$$b = \frac{8287.5}{17150.4} = 0.4832, \qquad a = 97.3 - 0.4832 \times 246.2 = -21.66$$

ゆえに, 回帰式: $y = -21.66 + 0.4832x$

(2) 散布図と回帰直線の描画



草量計による測定値と刈取草量の関係

(3) 分散分析（F 検定）

分散分析表は以下のとおりである．

H_0: 母分散比 $= 1$（回帰の効果はない）, H_1: 母分散比 > 1（回帰の効果がある）を立て, 上側の片側検定

回帰分析（$y = a + bx$）の分散分析表（練習問題 6-2）

変動因	平方和	自由度	分散	分散比（F）	有意性
回帰	36042.4	1	36042.4	267.1	<0.01
残差	1079.7	8	135.0	—	—
全体	37122.1	9	—	—	—

で有意水準 0.01 を採用すると, 付表 5 より $F(1, 8, 0.01) = 11.259$ が得られ, 棄却域は $F > 11.259$ となる. $F(= 267.1)$ は有意点より大きく棄却域に含まれるため, H_0 を棄却し, H_1 を採択する. すなわち, 有意水準 0.01 で回帰の効果があるといえる.

(4) 決定係数 R^2 と残差標準偏差 s_{e} の計算

決定係数 R^2 と残差標準偏差 s_{e} は以下のようになる.

$$R^2 = 36042.4/37122.1 = 0.971, \qquad s_{\mathrm{e}} = \sqrt{135.0} = 11.6$$

(5) 回帰式による推定

決定係数が 0.9 を超え, 回帰式の推定精度が高いので, 回帰式を使って, 草量計による測定値が 350 の地点の刈取草量を推定する.

$$y = -21.66 + 0.4832 \times 350 = 147.46$$

ゆえに, 刈取草量は約 $147.5\,\mathrm{g}$ 乾物/$0.25\,\mathrm{m}^2$ と推定される.

第 7 章

【7-1】

(1) 平方根変換

変換値（$\sqrt{x+1}$）

2.45	1.00	1.00	2.00	3.61	1.00	1.00	2.24	1.41	1.00
3.00	4.00	1.00	1.00	1.41	4.12	1.00	1.00	1.00	2.00

変換値の平均値 $= 1.81$, 平均値の逆変換値 $= 2.28$

(2) 対数変換

変換値（$\log x$）

1.74	2.88	3.69	2.31	3.27	2.36	2.11	2.43	2.23	1.43
2.36	2.33	3.08	3.18	2.39	2.76	2.98	1.70	1.85	1.65

変換値の平均値 $= 2.44$, 平均値の逆変換値 $= 273.1$

(3) 逆正弦変換

変換値（$\sin^{-1}\sqrt{x}$）

41.0	58.7	28.0	39.2	47.9	44.4	32.6	22.8	19.4	24.4
25.8	28.7	35.7	17.5	54.9	25.1	22.0	33.2	30.7	20.3

変換値の平均値 $= 32.6$, 平均値の逆変換値 $= 29.0$

(4) 逆正弦変換

変換値（$\sin^{-1}\sqrt{x}$）

25.0	33.5	27.3	29.8	34.5	25.8	26.0	8.5	31.4	29.0
33.0	30.9	19.4	30.3	14.9	26.6	23.7	27.1	32.5	18.8

変換値の平均値 $= 26.4$, 平均値の逆変換値 $= 19.8$