4-1

練習問題の答えなど

練習問題 4-2.1 の答え

(1) $1.581 \simeq 1.6$　(2) 1.0　(3) 0.50　(4) 0.25　(5) どんどん小さくなっていきます.

練習問題 4-2.2 の答え

	信頼度 90%	信頼度 95%	信頼度 99%
z_0 の値	1.64	1.96	2.58
信頼区間	$\mu = \bar{x} \pm 1.64\sigma/\sqrt{n}$	$\mu = \bar{x} \pm 1.96\sigma/\sqrt{n}$	$\mu = \bar{x} \pm 2.58\sigma/\sqrt{n}$

練習問題 4-2.3 の答え

(1)　信頼度 90%

$$\mu = \bar{x} \pm 1.64\sigma/\sqrt{n} \\ = 158.2 \pm 3.3$$

(2)　信頼度 95%

$$\mu = \bar{x} \pm 1.96\sigma/\sqrt{n} \\ = 158.2 \pm 3.9$$

(3)　信頼度 99%

$$\mu = \bar{x} \pm 2.58\sigma/\sqrt{n} \\ = 158.2 \pm 5.2$$

練習問題 4-2.4 の答え

この問題で与えられている数値 $\bar{x} = 25.6\,\mathrm{km/l}$, $\sigma = 2.0\,\mathrm{km/l}$, $n = 16$ を用います. 標準誤差 $\sigma/\sqrt{n} = 0.50$ なので, 答えは次のようになります.

(1)　信頼度 80 %：$\mu = \bar{x} \pm 1.28\sigma/\sqrt{n} = 25.6 \pm 0.64$　または,　$25.0 \leq \mu \leq 26.2$

(2)　信頼度 90 %：$\mu = \bar{x} \pm 1.64\sigma/\sqrt{n} = 25.6 \pm 0.82$　または,　$24.8 \leq \mu \leq 26.4$

(3)　信頼度 95 %：$\mu = \bar{x} \pm 1.96\sigma/\sqrt{n} = 25.6 \pm 0.98$　または,　$24.4 \leq \mu \leq 26.8$

練習問題 4-3.1 の答え

(1)　n=4 のとき, $P(t_0 \leq t) = 5\%$　$\rightarrow$　$t_0 = [\quad 2.353 \quad]$

(2)　n=4 のとき, $P(-t_0 \leq t \leq t_0) = 90\%$　$\rightarrow$　$t_0 = [\quad 2.353 \quad]$

(3)　n=9 のとき, $P(-t_0 \leq t \leq t_0) = 90\%$　$\rightarrow$　$t_0 = [\quad 1.860 \quad]$

(4)　n=9 のとき, $P(-t_0 \leq t \leq t_0) = 95\%$　$\rightarrow$　$t_0 = [\quad 2.306 \quad]$

(5)　n=16 のとき, $P(-t_0 \leq t \leq t_0) = 99\%$　$\rightarrow$　$t_0 = [\quad 2.947 \quad]$

練習問題 4-3.2 の答え

(1) $\bar{x} = 60.0, \quad s^2 = 36/8 = 9/2 = 4.50, \quad s = 3/\sqrt{2} = 2.12$. ついでに，$s/\sqrt{n} = 1/\sqrt{2} = 0.707$.

(2) 信頼度 95%なので，

$$P(-t_0 \leq t \leq t_0) = 0.95 \Leftrightarrow P(t_0 \leq t) = 0.025.$$

t 分布表を用いると，$n = 9$, $P(t_0 \leq t) = 0.025$ の時は，$t_0 = 2.306$.
したがって，信頼度 95% では，

$$-2.306 \leq t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq 2.306$$

これを解いて，

$$\bar{x} - 2.306 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 2.306 \cdot s/\sqrt{n}$$

となります．これに (1) で求めた数値を代入して，

$$60.0 - 1.630 \leq \mu \leq 60.0 + 1.630$$

$$58.4 \leq \mu \leq 61.6$$

つまり，信頼度 95%で，信頼区間は $58.4 \leq \mu \leq 61.6$（単位: グラム）となります．

練習問題 4-3.3 の答え

(1) $\bar{x} = 37.50, \quad s^2 = 0.36/8 = 0.4500, \quad s = 0.3/\sqrt{2} = 0.2121$. ついでに，$s/\sqrt{n} = 0.1/\sqrt{2} = 0.07071$.
（s, $s/\sqrt{n}$ は前問 (1) の 1/10 です.）

(2) 信頼度 95% では，

$$-2.306 \leq t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq 2.306$$

これを解いて，

$$\bar{x} - 2.306 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 2.306 \cdot s/\sqrt{n}$$

となります．これに (1) で求めた数値を代入して，

$$37.50 - 0.1630 \leq \mu \leq 37.50 + 0.1630$$

$$37.34 \leq \mu \leq 37.66$$

つまり，信頼度 95%で，信頼区間は $37.34 \leq \mu \leq 37.66$ (単位 : 度）となります．
というわけで，学校を休むかどうかは微妙なところ．

練習問題 4-3.4 の答え

今，$\bar{x} = 300$ 日，$s = 25$ 日，$n = 100$ なので，$s/\sqrt{n} = 2.5$ 日.

(1) $n = 100$, $P(-t_0 \leq t \leq t_0) = 80\%$ なので, 表から $t_0 = 1.290$. したがって，信頼度 80% では,

$$-1.290 \leq t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq 1.290$$

これを解いて,

$$\bar{x} - 1.290 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 1.290 \cdot s/\sqrt{n}$$

となります．この式に数値（$\bar{x} = 300$ 日，$s/\sqrt{n} = 2.5$ 日）を代入して，$296.8 \leq \mu \leq 303.2$ 日となります.

(2) $n = 100$, $P(-t_0 \leq t \leq t_0) = 90\%$ なので, 表から $t_0 = 1.660$. したがって,

$$\bar{x} - 1.660 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 1.660 \cdot s/\sqrt{n}$$

となります．この式に数値を代入して，$295.9 \leq \mu \leq 304.2$ 日となります.

(3) $n = 100$, $P(-t_0 \leq t \leq t_0) = 95\%$ なので, 表から $t_0 = 1.984$. したがって,

$$\bar{x} - 1.984 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 1.984 \cdot s/\sqrt{n}$$

となります．この式に数値を代入して，$295.0 \leq \mu \leq 305.0$ 日となります.

(4) $n = 100$, $P(-t_0 \leq t \leq t_0) = 99\%$ なので, 表から $t_0 = 2.626$. したがって,

$$\bar{x} - 2.626 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 2.626 \cdot s/\sqrt{n}$$

となります．この式に数値を代入して，$293.4 \leq \mu \leq 306.6$ 日となります.

練習問題 4-3.5 の答え

この問題では，標本の大きさが非常に大きいため t 分布表の $n = \infty$ の欄を使います．このように，大標本の場合は正規分布表を使えば良いのです.

今，$\bar{x} = 151.6$ cm，$s = 5.2$ cm，$n = 5000$ なので，$s/\sqrt{n} = 0.074$.

(1) 信頼度 95%のとき，表から $t_0 = 1.960$ なので,

$$\bar{x} - 1.960 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 1.960 \cdot s/\sqrt{n}$$

となります．数値（$\bar{x} = 151.6$ cm，$s/\sqrt{n} = 0.074$）を代入して $151.5 \leq \mu \leq 151.7$ cm となります.

(2) 信頼度 99%のとき，表から $t_0 = 2.576$ なので,

$$\bar{x} - 2.576 \cdot s/\sqrt{n} \leq \mu \leq \bar{x} + 2.576 \cdot s/\sqrt{n}$$

となります．この式に数値を代入して，$151.4 \leq \mu \leq 151.8$ cm となります.

(3) 標本の大きさを n 人すると，信頼度 99% の時の誤差は，$2.576 \cdot 5.2/\sqrt{n}$ となります．これが 0.1cm 以下になれば良いので，次のような不等式が成り立ちます.

$$2.576 \cdot 5.2/\sqrt{n} \leq 0.1.$$

これを解いて，$n \geq 17943 \simeq 18000$. したがって，約 18000 人以上ならば誤差が 0.1cm 以下になります.

練習問題 4-4.1 の答え

$\bar{p} = 6/10 = 0.6 (= 60\%)$ したがって，標本世帯の視聴率は 60% になります．

練習問題 4-4.2 の答え

視聴率調査は標本調査です．仮に全世帯に調査をして視聴率を求めたら，それが「全体の視聴率」です．そして，全体の視聴率 p と標本調査によって得られた視聴率 $\bar{p}$ との差が「誤差」です．

(1) $\bar{p} = 240/600 = 0.400 (= 40.0\%)$. したがって，標本世帯の視聴率は 40%になります．

(2) 関東では，600 世帯を調査したのですから，$n = 600$ ですね．標本世帯の視聴率を $\bar{p}$ だとすると，視聴率の標準誤差は，次のようになります．

$$S_E = \sqrt{\frac{\bar{p}(1-\bar{p})}{600}}$$

今，$\bar{p} = 0.400$ なので，

$$S_E = \sqrt{\frac{0.400 \cdot 0.600}{600}} = \sqrt{\frac{0.0400}{100}} = \frac{0.200}{10} = 0.0200 (= 2.00\%)$$

となります．

(3) 信頼度 68.26% の時の誤差を標準誤差と言うので，信頼度は 68.26%．また，全体の視聴率の信頼区間は

$$p = 0.400 \pm 0.0200 \quad \Leftrightarrow \quad 0.380 \le p \le 0.420 \quad (38.0\% \le p \le 42.0\%)$$

となります．

練習問題 4-4.3 の答え

標本世帯の視聴率 (標本平均) は $\bar{p} = 0.10 \quad (= 10\%)$，標準誤差は $\sqrt{0.10 \cdot 0.90/100} = 0.030 \quad (= 3.0\%)$ です．信頼度 90%, 95%, 99% の時の Z スコアの値は，標準正規分布表を用いて，それぞれ $z_0 = 1.645, z_0 = 1.960, z_0 = 2.576$ と求められます．

したがって，信頼度 90% では，

$$p = 0.10 \pm 1.645 \cdot 0.030 = 0.10 \pm 0.049 \quad \Leftrightarrow \quad 0.051 \le p \le 0.149 \quad (5.1\% \le p \le 14.9\%)$$

となります．また，信頼度 95% では，

$$p = 0.10 \pm 1.960 \cdot 0.030 = 0.10 \pm 0.059 \quad \Leftrightarrow \quad 0.041 \le p \le 0.159 \quad (4.1\% \le p \le 15.9\%)$$

となります．そして，信頼度 99% では，

$$p = 0.10 \pm 2.576 \cdot 0.030 = 0.10 \pm 0.077 \quad \Leftrightarrow \quad 0.023 \le p \le 0.177 \quad (2.3\% \le p \le 17.7\%)$$

となります．

練習問題 4-4.4 の答え

標本平均と標準誤差はそれぞれ次のようになります．

$$\bar{p} = \frac{1600}{2000} = 0.8000 \quad (= 80.00\%),$$
$$S_E = \sqrt{\frac{0.8 \cdot 0.2}{2000}} = 0.008944 \quad (= 0.8944\%)$$

信頼度 95% では，母比率 p の信頼区間は次のようになります．

$$p = \bar{p} \pm 1.960 \cdot S_E = 0.8000 \pm 0.01753 \quad \Leftrightarrow \quad 0.7825 \le p \le 0.8175. \quad (78.25\% \le p \le 81.75\%)$$

練習問題 4-4.5 の答え

標本検査の結果で得られた免疫を持ってない学生の割合 $\bar{p}$ とその標準誤差 S_E は次のようになります.

$$\bar{p} = \frac{10.0}{100} = 0.100 \quad (= 10.0\%),$$
$$S_E = \sqrt{\frac{0.100 \cdot 0.900}{100}} = 0.0300 (= 3.00\%)$$

信頼度 95% では免疫を持っていない学生の割合は次のようになると推測されます.

$$p = 0.100 \pm 1.960 \cdot 0.0300 = 0.100 \pm 0.0588 \quad \Leftrightarrow \quad 0.041 \leq p \leq 0.159 \quad (4.1\% \leq p \leq 15.9\%)$$

免疫を持っていない学生の人数を N 人とすると，$N = 4000 \cdot p$ なので，

$$4000 \cdot 0.041 \leq N \leq 4000 \cdot 0.159,$$

$$164\text{ 人} \leq N \leq 636\text{ 人}$$

となります.

また，ワクチンの費用は 1 人当たり 5000 円なので，全体でかかる費用は $5000N$ となります．つまり，

$$82.0\text{ 万円} \leq 5000N \leq 318\text{ 万円}$$

となります.

このように，標本 100 人程度では誤差がたいへん大きくなってしまいます．誤差を少なくするためには，標本の数を多く取れば良いのです．もし標本を 4 倍 (400 人) にしたとすると，誤差は半分になります.

練習問題 4-4.6 の答え

(1) (a) 視聴率 5%　　(b) 視聴率 10%　　(c) 視聴率 20%

信頼度が 95 % なので，次のようになります.

	(a)	(b)	(c)
$\bar{p}$	0.05	0.10	0.20
S_E	0.0089	0.012	0.016
$1.960 \cdot S_E$	0.017	0.024	0.032
真の視聴率	$0.033 \leq p \leq 0.067$	$0.076 \leq p \leq 0.124$	$0.168 \leq p \leq 0.232$
	$(3.3\% \leq p \leq 6.7\%)$	$(7.6\% \leq p \leq 12.4\%)$	$(16.8\% \leq p \leq 23.2\%)$

(2) 600 世帯のうち 6 世帯を買収してそのテレビ番組を見てもらったら，10% だった視聴率が 1% 上がって 11% になります．さて，(1) で求めたように，信頼度が 95% の時，真の視聴率の信頼区間は $10\% \pm 2.4\%$ です．つまり，買収して視聴率が増えましたが，統計的な誤差の範囲内です．したがって，6 世帯程度を買収しても統計的には有意な差は無い（意味が無い）と言えます.

(3) いいえ.
統計学を用いると，ある信頼度の時にどの程度の誤差になるかを理論的に推定できます．サンプル数 600 世帯というのは，全体からみれば確かに少ないかもしれませんが，中心極限定理が成り立つ十分なサンプル数であると考えられます．従って，たとえサンプル数が全体に対して少なかったとしても，得られた標本が無作為抽出（ランダムサンプリング）されたものであるならば，その推定値は正しいのです．もしもっと多くの標本を取ったら，統計的な誤差が減り，より精度の高い推定値を得ることができるというだけです.

練習問題 4-4.7 の答え

(1) $\bar{p}=0.40$ なので，標準誤差は $S_E=\sqrt{\frac{0.40\cdot 0.60}{n}}$ です．信頼度 90% なので，誤差は $1.645\cdot S_E = 1.645\cdot\sqrt{\frac{0.24}{n}}$ です．これが 0.01　(=1%) 以下なので次の式が成り立ちます．

$$1.645\sqrt{\frac{0.24}{n}} \leq 0.01$$

これを解くと，

$$n \geq 6494 \text{ 人}$$

となります．

(2) $\bar{p}=0.50$ なので，標準誤差は $S_E=\sqrt{\frac{0.50\cdot 0.50}{n}}$ です．信頼度 90% なので，誤差は $1.645\cdot S_E = 1.645\cdot\sqrt{\frac{0.25}{n}}$ です．これが 0.01　(=1%) 以下なので次の式が成り立ちます．

$$1.645\sqrt{\frac{0.25}{n}} \leq 0.01$$

これを解くと，

$$n \geq 6765 \text{ 人}$$

となります．

(3) $\bar{p}=0.60$ なので，標準誤差は $S_E=\sqrt{\frac{0.60\cdot 0.40}{n}}$ です．以降，(1) と同じで，$n \geq 6494$ 人 となります．

(4) $\bar{p}=0.70$ なので，標準誤差は $S_E=\sqrt{\frac{0.70\cdot 0.30}{n}}$ です．信頼度 90% なので，誤差は $1.645\cdot S_E = 1.645\cdot\sqrt{\frac{0.21}{n}}$ です．これが 0.01　(=1%) 以下なので次の式が成り立ちます．

$$1.645\sqrt{\frac{0.21}{n}} \leq 0.01$$

これを解くと，

$$n \geq 5683 \text{ 人}$$

となります．

(5) 以上の考察により，$\bar{p}=50\%$ のときに誤差 $1.645\cdot S_E$ は最大になります．(もちろんこれは数学的に厳密に証明する事もできますが，省略します．) したがって，この時に誤差が 1% 以下ならば，他の時にも 1% 以下になります．したがって，最低 6765 人に調査をすれば，すべての場合について 1% 以下になります．