練習問題（6章から9章の一部抜粋）

1. 100 人中に 1 月 1 日生まれの人がまったくいない確率を二項分布から求めよ。また、ポアソン分布による近似と比較せよ。
2. $X \sim U[0, 10]$ のとき、 $Y = [X]$ の確率分布を求めよ。ただし、 $[a]$ は実数 a を越えない最大の整数を表す。
3. 正規母集団からサイズ 9 のランダム標本を抽出するとき、標本平均と母集団平均の差の絶対値が母集団の標準偏差の半分を超える確率を求めよ。
4. 7-6 節で述べた大数の法則を、確率変数に対するチェビシェフの不等式を使って証明せよ。
5. 成功の確率 p の独立試行において、試行回数を n 、成功する回数を X とするとき、
(i) $\text{plim } X/n = p$ (ii) $\text{plim}(X-np) = 0$
を示せ。
6. 正規母集団からのランダム標本として、次のデータを得た。母集団平均、分散、標準偏差の 95% 信頼区間を構成せよ。
62, 58, 44, 42, 55, 58, 45, 64, 40, 52 (計 10 個)
7. 帰無仮説が受容されたとしても、必ずしも帰無仮説が正しいことにはならないことを、検定の論理から説明せよ。
8. 平均に関する左片側検定において、有意水準を小さくすると第 2 種の誤りの確率が高くなる関係があることを、帰無分布と対立分布を描くことによって確かめよ。
9. コインを 3 回投げて、3 回とも表あるいは裏が出たら、コインにゆがみがないという仮説を棄却するとき、第 1 種の誤りの確率を求めよ。また、この場合の検出力の表現を求めよ。
10. 1991 年の家計調査によれば、約 8,000 世帯の収入の平均は 676 万円、標準偏差は 410 万円である。母集団世帯の収入の平均は 670 万円よりも多いと考えられるか。
11. 入学試験のある科目は満点が 120 点であり、採点者達は得点の標準偏差が少なくとも 15 点になるような採点を意図した。実際の採点においてランダムに 25 人の点数を調べたところ、標準偏差は 17 点であった。採点者達の意図は達成されていると考えられるか。

練習問題解答

1. $X \sim B(100, 1/365)$ とするとき,

$$P(X=0) = (364/365)^{100} \approx 0.76007$$

ポアソン近似では, $\lambda = np = 100/365$ として,

$$\exp(-100/365) \approx 0.76035$$

2. $P(X=y) = P(X < y+1)$ を使う.

$$P(Y=y) = P(X=y) = P(Y < y+1) = 1/10 \quad (y=0, 1, \dots, 9)$$

すなわち, Y は $[0, 9]$ 上の離散的一様分布に従う.

$$3. P(|\bar{X} - \mu| > \sigma/2) = P(\sqrt{n} |\bar{X} - \mu| / \sigma > \sqrt{n}/2) = P(|Z| > 1.5) = 0.1336$$

$$4. P(|\bar{X} - \mu| > \varepsilon) = P((\bar{X} - \mu)^2 < \varepsilon^2) < E[(\bar{X} - \mu)^2] / \varepsilon^2 = \sigma^2 / n\varepsilon^2 \rightarrow 0$$

$$5. (i) P(|X/n - p| > \varepsilon) < E[(X/n - p)^2] / \varepsilon^2 = p(1-p) / n\varepsilon^2 \rightarrow 0$$

$$(ii) P(|X - np| > \varepsilon) > P(X > np + \varepsilon)$$

$$= P\left(\frac{X - np}{\sqrt{np(1-p)}} > \frac{\varepsilon}{\sqrt{np(1-p)}}\right) \rightarrow P(Z > 0) = \frac{1}{2}$$

$$6. \text{平均の信頼区間: } \bar{X} - t_{\alpha/2}^{(n-1)} \frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2}^{(n-1)} \frac{\hat{\sigma}}{\sqrt{n}}$$

$$\Rightarrow 52 - 2.262 \times \frac{8.705}{\sqrt{10}} \leq \mu \leq 52 + 2.262 \times \frac{8.705}{\sqrt{10}}$$

$$\text{分散の信頼区間: } \frac{(n-1)\hat{\sigma}^2}{c_{\alpha/2}^{(n-1)}} \leq \sigma^2 \leq \frac{(n-1)\hat{\sigma}^2}{c_{1-\alpha/2}^{(n-1)}}$$

$$\Rightarrow \frac{9 \times 75.7778}{19.022} \leq \sigma^2 \leq \frac{9 \times 75.7778}{2.700}$$

$$\text{標準偏差の信頼区間: } \sqrt{\frac{(n-1)\hat{\sigma}^2}{c_{\alpha/2}^{(n-1)}}} \leq \sigma \leq \sqrt{\frac{(n-1)\hat{\sigma}^2}{c_{1-\alpha/2}^{(n-1)}}}$$

7. 論理の問題として, A (帰無仮説) が正しいと仮定して, A に矛盾しない結果が得られたとしても, 必ずしも A が正しいことにはならない. A に矛盾する結果が得られたならば, A は正しくないといえる.

8. 右片側検定の場合については，テキスト 193 ページの図 9-1 にあるので，これを左片側検定の場合に応用すればよい．

9. X を表が出る回数とするとき，

$$\text{第 1 種の誤りの確率} = P(\{X = 0\} \cup \{X = 3\} | H_0) = 2 \times \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

$$\text{検出力} = P(\{X = 0\} \cup \{X = 3\} | H_1) = (1-p)^3 + p^3$$

10. 検定問題 $H_0: \mu = 670$ vs $H_1: \mu > 670$ を考える．

$$\text{検定統計値} = \frac{\sqrt{8000}(676 - 670)}{410} = 1.31 < z_{0.05} = 1.645$$

$\therefore$ 670 万円よりも多いとは考えにくい．

11. 検定問題 $H_0: \sigma = 15$ vs $H_1: \sigma > 15$ を考える．

$$\text{検定統計値} = \frac{24 \times 17^2}{15^2} = 30.83 < c_{0.05}^{(24)} = 36.42$$

$\therefore$ 意図が達成されているとはいえない．