練習問題2 解答

(答えに関して質問等ある場合は nagakura@z7.keio.jp まで連絡して下さい)

(1) 答 (F)

$\Pr(-2 \le X \le 2)$は平均を引き標準偏差で割ることによって

$$\Pr\left((-2-2)/2 \le (X-2)/2 \le (2-2)/2 \right) = \Pr(-2 \le Z \le 0)$$

と書き換えることができる。ここで $Z = (X-2)/2$ である。Z は標準正規分布に従うので分布表より、$\Pr(-2 \le Z \le 0)$となる確率を求めればよい(標準正規分布は 0 で対称な分布なので

$$\Pr(-2 \le Z \le 0) = \Pr(0 \le Z \le 2)$$
$$= \Pr(Z \le 2) - 0.5$$

を求めてもよい)。分布表より $\Pr(Z \le 2) - 0.5 \approx 0.4772$ なので答えは (F) になる。

(2) 答 (H)

標本平均は 40/50 = 0.8 であるので答えは (H)である。

(3) 答 (F)

$\mathrm{var}(X) = 1$, $\mathrm{var}(Y) = 4$ であるので

$$\mathrm{corr}(X, Z) = \frac{\mathrm{cov}(X, Z)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Z)}} = \frac{\mathrm{cov}(X,\ 3X + 2Y)}{\sqrt{\mathrm{var}(3X + 2Y)}}$$
$$= \frac{3\,\mathrm{cov}(X, X)}{\sqrt{9\,\mathrm{var}(X) + 4\,\mathrm{var}(Y)}}$$
$$= \frac{3}{\sqrt{25}} = \frac{3}{5} = 0.6.$$

となる。よって答は (F) となる。

(4) 答 (I)

$X \sim N(2, 16)$ より $Z = (X-2)/4$ は標準正規分布に従う。標準正規分布の分布表より

$$\Pr(Z \le 0.675) \approx 0.75$$

である。よって

$$\Pr\left((X-2)/4 \le 0.675 \right) = \Pr(X \le 4.7)$$

であるので、c の値はおよそ 4.7 である事がわかる

(5) 答 1/3

$\Pr(3Y + 2 \le 3) = \Pr(Y \le 1/3) = 1/3$ ($U(0, 1)$ の性質より)

(6) 答 (B)

正規分布の和は正規分布である。ただし、平均と分散は変わるので注意。

(7) 答 1/3

$U(a, b)$の分散は$(b-a)^2/12$ なので

$$\mathrm{var}(X) = \frac{2^2}{12} = \frac{4}{12} = \frac{1}{3}$$

となる。

(8) 答 $\exp(-y)$

$$\Pr(Y \le y) = \Pr(-\log X \le y) = \Pr(X \ge \exp(-y)) = 1 - \Pr(X \le \exp(-y)) = 1 - \exp(-y)$$

(最後の等式は一様分布の性質より)。これより Y の密度関数は

$$\frac{d\,\Pr(Y \le y)}{dy} = \frac{1 - \exp(-y)}{dy} = \exp(-y)$$

となる。

(9) 答 $-\sigma^2/n$

$$E(\tilde{s}^2) = \frac{n-1}{n} E\left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right) = \frac{n-1}{n}\sigma^2$$

であるので(標本分散は σ^2 の不偏推定量である事に注意)、$\tilde{s}^2$ のバイアスは

$$E(\tilde{s}^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \frac{n-1-n}{n}\sigma^2 = -\frac{1}{n}\sigma^2$$

となる。

(10) 答 μ^2

$$E(\tilde{\sigma}^2) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i^2 \right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) = \frac{1}{n} n(\sigma^2 + \mu^2) = \sigma^2 + \mu^2$$

であるので $\tilde{\sigma}^2$ のバイアスは

$$E(\tilde{\sigma}^2) - \sigma^2 = \sigma^2 + \mu^2 - \sigma^2 = \mu^2$$

となる。

(11) 答 $3x^2$

$U(0, 1)$からの大きさ 3 の i.i.d. 標本を $\{U_1, U_2, U_3\}$ とする。確率変数 X は$\{U_1, U_2, U_3\}$の 3 つの最大値であるので X の分布関数は

$$F_x(x) = \Pr(X \le x) = \Pr(U_1 \le x) \Pr(U_2 \le x) \Pr(U_3 \le x)$$

$$= x^3 \quad (U(0, 1) \text{ の性質より } \Pr(U_j \le x) = x,\ j = 1,\ldots,3 \text{ なので})$$

となる。よってその密度関数は $f_x(x) = \mathrm{d}F_x(x)/\mathrm{d}x = 3x^2$ となる。同様の議論により、一般に $U(0, 1)$からの大きさ m の i.i.d. 標本の最大値の分布は $F_x(x) = x^m$ となる。よってその密度関数は $f_x(x) = mx^{m-1}$ となる。

(12) 答 3/4

X の密度関数が $f_x(x) = 3x^2$ であるので、$E(X)$は

$$E(X) = \int_0^1 x\, f_x(x)dx = 3\int_0^1 x^3\, dx = 3\left[\frac{1}{4}x^4\right]_0^1 = \frac{3}{4}$$

となる。

(13) 答 4%

有意水準とは帰無仮説を誤って棄却する確率である。有意水準を $100\alpha\,\%$ とすると、この問題の場合は

$$\alpha = \int_{\log(25)}^{\infty} \exp(-t)dt$$

となる。これを計算すると

$$\begin{aligned}\alpha = \int_{\log(25)}^{\infty} \exp(-t)dt &= \left[-\exp(-t)\right]_{\log(25)}^{\infty} \\ &= \exp[-\log(25)] \\ &= \exp[\log(1/25)] = 1/25 = 0.04\end{aligned}$$

となるので、有意水準は 4%である。

(14) 答 0.5196

p の推定値は、$\hat{p}_{2500} = \frac{1}{2500}\sum_{i=1}^{1250} X_i = \frac{1250}{2500} = 0.5$ となる。この推定量は標本平均なので、基準化し

たものは中心極限定理により、n が大きいとき標準正規分布に従う(でよく近似できる)。つまり

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sigma} \sim N(0,1),\ \ \sigma^2 = \mathrm{var}[X_i]$$

である(ここで母分散 σ^2 は $\sigma^2 = p(1-p)$である事に注意)。よって(講義スライドでの議論より)95%信頼区間は、σ をその推定値で置き換えて

$$\begin{aligned}&\left[\hat{p}_{2500} - 1.96\frac{\sqrt{\hat{p}_{2500}(1-\hat{p}_{2500})}}{\sqrt{n}},\ \hat{p}_{2500} + 1.96\frac{\sqrt{\hat{p}_{2500}(1-\hat{p}_{2500})}}{\sqrt{n}}\right] \\ &= \left[0.5 - 1.96\frac{\sqrt{0.25}}{\sqrt{2500}},\ 0.5 + 1.96\frac{\sqrt{0.25}}{\sqrt{2500}}\right] \\ &= \left[0.5 - 1.96\frac{0.5}{50},\ 0.5 + 1.96\frac{0.5}{50}\right] \\ &= [0.5 - 0.0196,\quad 0.5 + 0.0196] \\ &= [0.4804,\ 0.5196]\end{aligned}$$

となる。よってその上限は 0.5196 となる。

(15) 答: 有意水準 5%では棄却、1%では採択(棄却できない)となる。

問題(14)の時と同様に、p の標本平均による推定値は

$$\hat{p}_{2400} = \frac{1380}{2400} = 0.575$$

となる。また、検定統計量として、帰無仮説の下での値 p_0 を用いて標本平均を基準化した統計量

$$Z_n = \frac{\sqrt{n}(\hat{p}_n - p_0)}{\sqrt{p_0(1-p_0)}}$$

を用いて検定する。今、$n = 2400$, $\hat{p}_n = 0.575$, $p_0 = 0.6$ なので、Z_n の実現値は

$$Z_{2400} = \frac{\sqrt{2400}\times(0.575-0.6)}{\sqrt{0.6(1-0.6)}} = \frac{10\sqrt{24}\times(-0.025)}{\sqrt{0.24}}$$

$$= \sqrt{\frac{24}{0.24}}\times(-0.25) = \sqrt{100}\times(-0.25) = 10\times(-0.25) = -2.5$$

となる。対立仮説は H_1: $p \neq 0.6$ なのでこれは両側検定である。よって有意水準 5%、1%のとき、Z_{2400} の絶対値がそれぞれ 1.96 および 2.58 より大きい時に帰無仮説を棄却する。今、$|Z_{2400}| = 2.5$ なので、この場合有意水準 5%では棄却できるが、1%では棄却できない、となる。

(16) 答　有意水準 5%では棄却、1%でも棄却、となる。

帰無仮説、および標本数 n は問題(15)の時と同じであるので Z_n の実現値は問題(15)の時と同じである。即ち $Z_n = -2.5$。対立仮説は $p < 0.6$ であるのでこれは(下側)片側検定となる。よって Z_n が -1.64、-2.33 より小さい時に、それぞれ有意水準 5%、1%で帰無仮説を棄却する事になる。よって、5%でも 1%でも共に棄却、となる。

(17) 答 採択(棄却されない)、となる

中心極限定理により Z_n は標本数が大きい時、標準正規分布に従う。よって Z_n の絶対値が 1.96 より大きい場合に有意水準 5%で棄却される。$Z_n = -1.4$ であるので、答えは「棄却されない」となる。