問題解答（第 2 章）

解答 2.1　分散が最大となるのは，$V[R] = \theta(1-\theta)$ を θ で微分して 0 と置き，$1-2\theta = 0$ より $\theta = 0.5$ である．$\theta = 0.5$ のとき R は 0 と 1 を等確率で取るため最もばらつきが大きいと解釈される．

解答 2.2　(2.5) より $\mu'_m = E[R^m] = \theta$ であるので

$$\mu_3 = E[(R-\theta)^3] = E[R^3 - 3\theta R^2 + 3\theta^2 R - \theta^3] = \theta(1-\theta)(1-2\theta)$$

$$\mu_4 = E[(R-\theta)^4] = E[R^4 - 4\theta R^3 + 6\theta^2 R^2 - 4\theta^3 R + \theta^4] = \theta(1-\theta)(1-3\theta+3\theta^2)$$

を得る．よって，歪度と尖度は

$$\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{\theta(1-\theta)(1-2\theta)}{\theta(1-\theta)\sqrt{\theta(1-\theta)}} = \frac{1-2\theta}{\sqrt{\theta(1-\theta)}},$$

$$\beta_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\theta(1-\theta)(1-3\theta+3\theta^2)}{\theta^2(1-\theta)^2} - 3 = \frac{1-6\theta+6\theta^2}{\theta(1-\theta)}$$

となる．$\theta = 0.5$ のとき，分布は 0.5 を中心に左右対称であるので $\beta_1 = 0$ となる．またこのとき尖度は最小値 $\beta_2 = -2$ を取る．

解答 2.3　$\dfrac{p(x+1)}{p(x)} = \dfrac{n-x}{x+1} \cdot \dfrac{\theta}{1-\theta} \geq 1$ を変形した不等式 $(n-x)\theta \geq (x+1)(1-\theta)$ を θ について解いて $\dfrac{x+1}{n+1} \geq \theta$ を得る．逆向きの不等式も成り立つ．

解答 2.4　確率母関数 $H_X(t) = \{\theta t + (1-\theta)\}^n$ を t で 2 回微分すると

$$\frac{d}{dt}H_X(t) = n\theta\{\theta t + (1-\theta)\}^{n-1}, \quad \frac{d^2}{dt^2}H_X(t) = n(n-1)\theta^2\{\theta t + (1-\theta)\}^{n-2}$$

であるので，ここで $t = 1$ と置いて $E[X(X-1)] = n(n-1)\theta^2$ を得る．

解答 2.5　期待値は

$$E[X] = \frac{d}{dt}\{\theta e^t + (1-\theta)\}^n |_{t=0} = n\theta \exp(t)\{\theta e^t + (1-\theta)\}^{n-1} |_{t=0} = n\theta$$

と求められる．分散は

$$
\begin{aligned}
E[X^2] &= \frac{d^2}{dt^2}\{\theta e^t + (1-\theta)\}^n |_{t=0} \\
&= [n\theta e^t \{\theta \exp(t) + (1-\theta)\}^{n-1} + n(n-1)\theta^2 \{e^t\}^2 \{\theta e^t + (1-\theta)\}^{n-2}]|_{t=0} \\
&= n\theta + n(n-1)\theta^2
\end{aligned}
$$

より

$$V[X] = E[X^2] - (E[X])^2 = n\theta + n(n-1)\theta^2 - (n\theta)^2 = n\theta(1-\theta)$$

となる.

解答 2.6 $X \sim B(m,\ \theta)$, $Y \sim B(n,\ \theta)$ で X と Y は互いに独立とし, それらの和を $W = X + Y$ とする. このとき

$$
\begin{aligned}
\Pr(W = w) &= \sum_{x+y=w} {}_mC_x\, \theta^x (1-\theta)^{m-x}\, {}_nC_y\, \theta^y (1-\theta)^{n-y} \\
&= \sum_{x+y=w} {}_mC_x \times {}_nC_y\, \theta^{x+y} (1-\theta)^{(m+n)-(x+y)} = \theta^w (1-\theta)^{(m+n)-w} \sum_{x+y=w} {}_mC_x \times {}_nC_y
\end{aligned}
$$

である. $\sum_{x+y=w} {}_mC_x \times {}_nC_y = {}_{m+n}C_w$ がいえれば $\Pr(W = w) = {}_{m+n}C_w\, \theta^w (1-\theta)^{(m+n)-w}$ より $W \sim B(m + n,\ \theta)$ となる. それは, z を実数とした等式 $(1 + z)^{m+n} = (1 + z)^m(1 + z)^n$ の両辺の展開

$$\sum_{w=0}^{m+n} {}_{m+n}C_w\, z^w = \left(\sum_{x=0}^{m} {}_mC_x\, z^x\right)\left(\sum_{y=0}^{n} {}_nC_y\, z^y\right) = \sum_{w=0}^{m+n} \sum_{x+y=w} {}_mC_x \times {}_nC_y\, z^w$$

での z^w の係数の比較により示される.

解答 2.7 N 個の個体の集合からランダムに n 個選ぶ組み合わせは全部で ${}_NC_n$ 通りある. それらのうち, 特性 A をもつ M 個から y 個選ぶ組み合わせは ${}_MC_y$ 通りあり, 特性 A を持たない残りの $N - M$ 個から $n - y$ 個選ぶ組み合わせは ${}_{N-M}C_{n-y}$ 通りである. よって, M 個から y 個かつ $N - M$ 個から $n - y$ 個選ぶ場合の数は ${}_MC_y \times {}_{N-M}C_{n-y}$ となり, これを ${}_NC_n$ で割って確率 (2.13) を得る. y の取り得る最大値が n と M のうちで小さい方, すなわち $\min(n, M)$ となるのは自明である. y の最小値は 0 か, あるいは $N - M$ 個の特性 A を持たない個体をすべて取りつくしてしまった場合に必然的に取らなければならない特性 A を持つ個体数 $\{n - (N - M)\}$ であるので $\max(0, n - N + M)$ となる.

解答 2.8 超幾何分布 $H(n, M, N)$ の期待値は,

$$E[Y]=\sum_y y\frac{{}_M C_y \times {}_{N-M} C_{n-y}}{{}_N C_n}=\sum_y y\frac{\dfrac{M!}{y!(M-y)!}\times\dfrac{(N-M)!}{(n-y)!\{(N-M)-(n-y)\}!}}{\dfrac{N!}{n!(N-n)!}}$$

であるが，ここで $k=y-1$ とおくと

$$=\frac{nM}{N}\sum_k \frac{{}_{M-1}C_k \times {}_{(N-1)-(M-1)}C_{(n-1)-k}}{{}_{N-1}C_{n-1}}$$

となる．最後の和はパラメータ $N-1$，$M-1$，$n-1$ の超幾何分布の全確率であるので 1 になることより $E[Y]=nM/N$ が示される．

解答 2.9　相関係数は，$R=R[X_j,X_k]=-\sqrt{\dfrac{1/m}{1-1/m}\times\dfrac{1/m}{1-1/m}}=-\dfrac{1}{m-1}$ である．また，$|R|\le 0.2$ となるのは $|-1/(m-1)|\le 0.2$ より $m\ge 6$ の場合である．

解答 2.10　$B(20, 0.5)$ における確率 $p(x)=\Pr(X=x\mid\theta=0.5)$ および P-値（下側累積確率）と mid-P 値は右の表のようである．これより，P-値 ≤ 0.05 となる最大の x は 5 であり，mid-P 値 ≤ 0.05 となる最大の x は 6 である．

x	p(x)	P-value	mid-P
0	0.000	0.000	0.000
1	0.000	0.000	0.000
2	0.000	0.000	0.000
3	0.001	0.001	0.001
4	0.005	0.006	0.004
5	0.015	0.021	0.013
6	0.037	0.058	0.039
7	0.074	0.132	0.095
8	0.120	0.252	0.192
9	0.160	0.412	0.332
10	0.176	0.588	0.500

解答 2.11　$B(20, 0.5)$ を近似する正規分布は $N(10, 5)$ である．$X\sim B(20, 0.5)$ および $Y\sim N(10, 5)$ としたとき，たとえば下側累積確率 $\Pr(X\le 6)$ の近似は，連続修正がない場合には $\Pr(Y\le 6)$ = NORMDIST(6, 20, SQRT(5), 1) ≈ 0.037 となり，連続修正がある場合には $\Pr(Y\le 6.5)$ = NORMDIST(6.5, 20, SQRT(5), 1) ≈ 0.059 となる．$x=0\sim 10$ に対する各確率は右の表のようである．下側累積確率が 0.05 以下となる最大の整数は，修正がある場合には 5，ない場合には 6 と問題 2.6 の P-値および mid-P 値の場合と対応している．

x	修正あり	修正なし
0	0.000	0.000
1	0.000	0.000
2	0.000	0.000
3	0.002	0.001
4	0.007	0.004
5	0.022	0.013
6	0.059	0.037
7	0.132	0.090
8	0.251	0.186
9	0.412	0.327
10	0.588	0.500

解答 2.12　$n=16$，$x=4$ であるので，点推定値は成功率 $\hat\theta=4/16=1/4=0.25$ によって与えられ，標準誤差は $SE[\theta]=\sqrt{\dfrac{1}{4}\left(1-\dfrac{1}{4}\right)/16}=\dfrac{\sqrt{3}}{16}\approx 0.108$ となる．

解答 2.13　$X \sim B(n, \theta)$ のとき，$E[X] = n\theta$ であり，$V[X] = n\theta(1-\theta)$ であるので，$E[X^2] = V[X] + (E[X])^2 = n\theta(1-\theta) + n^2\theta^2$ である．よって，

$$E\left[\frac{X}{n}\left(1-\frac{X}{n}\right)\right] = E\left[\frac{X}{n}\right] - E\left[\frac{X^2}{n^2}\right] = \theta - \frac{\theta(1-\theta)}{n} - \theta^2 =$$
$$= \theta(1-\theta) - \frac{\theta(1-\theta)}{n} = \frac{n-1}{n}\theta(1-\theta)$$

を得る．

解答 2.14　$n = 16$，$x = 4$ であるので，$\tilde{\theta}_1 = \frac{4+1}{16+2} = \frac{5}{18} \approx 0.278$ および $\tilde{\theta}_2 = \frac{4+2}{16+4} = \frac{6}{20} = 0.3$ となる．

解答 2.15　$n = 16$，$x = 4$ であるので，信頼下限における F 分布の自由度は $l_1 = 2 \times (16-4+1) = 26$, $l_2 = 2 \times 4 = 8$ であり，$\alpha = 0.05$ とすると，

$$F_{26,8}(0.025) = \mathrm{FINV}(0.025, 26, 8) \approx 3.927$$

であるので，$\theta_L = 4/(4 + 13 \times 3.927) = 0.073$ となる．上限については $k_1 = 2 \times (4+1) = 10$, $k_2 = 2 \times 12 = 24$ であり，

$$F_{10,24}(0.025) = \mathrm{FINV}(0.025, 10, 24) \approx 2.640$$

であるので，$\theta_U = 5/(5 + 12/2.640) = 0.524$ を得る．

解答 2.16　$n = 16$，$x = 0$ の場合，信頼下限は明らかに 0 である．信頼上限は $(1-\theta_U)^{16} = 0.025$ より $\theta_U = 1 - \sqrt[16]{0.025} \approx 0.206$ となる．F 分布を用いた計算では，自由度が $k_1 = 2$，$k_2 = 32$ であるので，

$$F_{2,\,32}(0.025) = \mathrm{FINV}(0.025, 2, 32) \approx 4.149$$

となり，$\theta_U = 1/(1 + 16/4.149) = 0.206$ を得る．この場合の信頼区間は $[0, 0.206)$ という半開区間となる．事象が 1 度も観測されなくても $\theta = 0.2$ 程度の可能性はあることになる．

解答 2.17　$n = 16$，$x = 4$ では，$c = z(0.025) \approx 1.96$ とした場合

スコア法 (2.28)：(0.102, 0.495)，ワルド法 (2.29)：(0.038, 0.462)

となる．$c = z(0.025)$ を 2 と近似すると

スコア法 (2.28)：(0.100, 0.500)，ワルド法 (2.29)：(0.033, 0.467)

となる．$\tilde{\theta}_2 = (x + 2)/(n + 4) = 6/20 = 0.3$ に基づくワルド法による信頼区間は

$$\theta_L = \tilde{\theta}_2 - 2\sqrt{\tilde{\theta}_2(1-\tilde{\theta}_2)/n} = 0.095, \quad \theta_U = \tilde{\theta}_2 + 2\sqrt{\tilde{\theta}_2(1-\tilde{\theta}_2)/n} = 0.505$$

となる．

解答 2.18　$\hat{\theta} = x/n = 0$ であるので (2.28) より下限は $\theta_L = 0$ となり，上限は $\theta_U = c^2/(n + c^2)$ となる．$n = 16$ では $\theta_U = c^2/(16 + c^2)$ であり，$c = 2$ とすると $\theta_U = 4/(16 + 4) = 4/20 = 0.2$ を得る．

解答 2.19　事後分布は $Beta(6, 14)$ であり，事後モードは $\tilde{\theta}_{\mathrm{mode}} = 5/18 \approx 0.278$ となる．95%確率区間は

$$\mathrm{BETAINV}(0.025, 6, 14, 0, 1) \approx 0.126,$$
$$\mathrm{BETAINV}(0.975, 6, 14, 0, 1) \approx 0.512$$

より (0.126, 0.512) となる．

解答 2.20　$\omega = \theta/(1 - \theta)$ を θ について解き直すと，$\omega - \omega\theta = \theta$ より $(1 + \omega)\theta = \omega$ となり，$\theta = \omega/(1 + \omega)$ を得る．$\omega = e^{\phi}$ であるのでこれを代入し $\theta = e^{\phi}/(1 + e^{\phi}) = 1/(e^{-\phi} + 1)$ となる．

解答 2.21　オッズ ω の推定値は $\hat{\omega} = 4/(16-4) = 1/3$ となり，対数オッズ ϕ の推定値は $\hat{\phi} = \log\hat{\omega} = \log(1/3) \approx -1.099$ となる．$\hat{\phi}$ の分散の推定値は $1/4 + 1/12 = 1/3$ であるので，ϕ の近似的な 95%信頼区間は $-1.099 \pm 1.96\sqrt{1/3} = (-2.230, 0.033)$ となる．オッズ ω の 95%信頼区間は $(e^{-2.230}, e^{0.033}) \approx (0.108, 1.034)$ となる．例 2.16 の $n = 16$，$x = 12$ のときの信頼区間は (0.968, 9.302) であったが，$0.108 = 1/9.302$，$1.034 = 1/0.968$ と逆数の関係になっている．

解答 2.22　期待値 $E[Y]$ は

$$E[Y] = \frac{1}{1-(1-\theta)^n}\sum_{y=1}^{n} y \times {}_nC_y \theta^y (1-\theta)^{n-y} = \frac{1}{1-(1-\theta)^n}\sum_{x=0}^{n} x \times {}_nC_x \theta^x (1-\theta)^{n-x}$$
$$= \frac{E[X]}{1-(1-\theta)^n} = \frac{n\theta}{1-(1-\theta)^n}$$

により示される．分散については，

$$
\begin{aligned}
E[Y(Y-1)] &= \frac{1}{1-(1-\theta)^n}\sum_{y=1}^{n} y(y-1)\times {}_nC_y \theta^y (1-\theta)^{n-y} \\
&= \frac{1}{1-(1-\theta)^n}\sum_{x=2}^{n} x(x-1)\times {}_nC_x \theta^x (1-\theta)^{n-x} \\
&= \frac{1}{1-(1-\theta)^n} E[X(X-1)] = \frac{n(n-1)\theta^2}{1-(1-\theta)^n}
\end{aligned}
$$

であることから,

$$
\begin{aligned}
V[Y] &= E[Y^2]-(E[Y])^2 = E[Y(Y-1)]+E[Y]-(E[Y])^2 \\
&= \frac{n(n-1)\theta^2}{1-(1-\theta)^n} + \frac{n\theta}{1-(1-\theta)^n} - \left\{\frac{n\theta}{1-(1-\theta)^n}\right\}^2 \\
&= \frac{n\theta(1-\theta)}{1-(1-\theta)^n} + \frac{(n\theta)^2}{1-(1-\theta)^n} - \frac{(n\theta)^2}{\{1-(1-\theta)^n\}^2} \\
&= \frac{n\theta(1-\theta)}{1-(1-\theta)^n} - \frac{(n\theta)^2(1-\theta)^n}{\{1-(1-\theta)^n\}^2}
\end{aligned}
$$

と示される.

解答 2.23　$ZTB(n, \theta)$ における相続く y に関する確率の比は

$$
\frac{\Pr(Y=y+1)}{\Pr(Y=y)} = \frac{n-y}{y+1}\cdot\frac{\theta}{1-\theta}
$$

となる. ここで $y = 1$ とした不等式

$$
\frac{\Pr(Y=2)}{\Pr(Y=1)} = \frac{n-1}{2}\cdot\frac{\theta}{1-\theta} \le 1
$$

より関係式 $\theta \le 2/(n+1)$ を得る. このとき, すべての y で $p(y) \ge p(y+1)$ も示されるので, $\theta \le 2/(n+1)$ が, 確率が y に関し単調減少となるための条件となる.

解答 2.24　方程式 $n\theta = 1-(1-\theta)^n$ は明らかに $\theta = 0$ の解を持つ. これが唯一の解であることは, 左辺と右辺の差を $f(\theta) = n\theta - \{1-(1-\theta)^n\}$ とすると, $f'(\theta) = n\{1-(1-\theta)^{n-1}\}$ であるので, $0 < \theta \le 1$ の範囲で $f'(\theta) > 0$ となり, $f(\theta)$ は単調増加関数であることから示される.

解答 2.25　$ZIB(n, \theta, \omega)$ におけるゼロ度数の確率は $\Pr(Y=0) = \omega + (1-\omega)(1-\theta)^n$ であり, $\Pr(Y=0) \ge 0$ より $\omega + (1-\omega)(1-\theta)^n \ge 0 \Rightarrow \omega\{1-(1-\theta)^n\} + (1-\theta)^n \ge 0$

$\Rightarrow\ \omega \geq -\dfrac{(1-\theta)^n}{1-(1-\theta)^n}$ となる．

解答 2.26　$ZIB(n,\ \theta,\ \omega)$ の分散 $V[Y] = (1-\omega)n\theta(1-\theta) + \omega(1-\omega)(n\theta)^2$ を ω で微分して 0 と置くと $-n\theta(1-\theta) + (1-2\omega)(n\theta)^2 = 0$ であるので，これより分散の最大値を与える ω として $\omega = \dfrac{1}{2}\left(1-\dfrac{1-\theta}{n\theta}\right)$ を得る．ω が下限 $-\dfrac{(1-\theta)^n}{1-(1-\theta)^n}$ に等しい場合は，

$$1-\omega = 1+\frac{(1-\theta)^n}{1-(1-\theta)^n} = \frac{1}{1-(1-\theta)^n}$$

であるので，$ZIB(n,\ \theta,\ \omega)$ の期待値および分散は

$$E[Y] = (1-\omega)n\theta = \frac{n\theta}{1-(1-\theta)^n},$$

$$V[Y] = (1-\omega)n\theta(1-\theta) + \omega(1-\omega)(n\theta)^2 = \frac{n\theta(1-\theta)}{1-(1-\theta)^n} - \frac{n^2\theta^2(1-\theta)^n}{\{1-(1-\theta)^n\}^2}$$

とゼロトランケートされた二項分布の期待値と分散に一致する．

解答 2.27　初期値 $M^{(0)} = N = 20$ から出発すると，右のような推移をたどり $\hat{\theta} = 0.5678$，$\hat{\omega} = 0.1370$，$\hat{M} = 17.2604$ に収束する．θ の値が大きいので収束は例 2.20 に比べ速い．

ITE	M	Theta	Omega
0	20.0000	0.4900	0.0000
1	17.6075	0.5566	0.1196
2	17.2965	0.5666	0.1352
3	17.2640	0.5677	0.1368
4	17.2607	0.5678	0.1370
5	17.2604	0.5678	0.1370
6	17.2604	0.5678	0.1370