問題解答（第 3 章）

解答 3.1　オッズの定義 $\omega = \theta_2/(1-\theta_2)$ より $1-\theta_2 = 1/(1+\omega)$ となる．また $\omega\psi = \theta_1/(1-\theta_1)$ から $1-\theta_1 = 1/(1+\omega\psi)$ を得る．$y = s-x$ であるので (3.6) より

$$
\begin{aligned}
&{}_mC_x \cdot {}_nC_{s-x} \cdot \left(\frac{\theta_1}{1-\theta_1}\right)^x \left(\frac{\theta_2}{1-\theta_2}\right)^{s-x} (1-\theta_1)^m (1-\theta_2)^n \\
&= {}_mC_x \cdot {}_nC_{s-x} \cdot \left\{\left(\frac{\theta_1}{1-\theta_1}\right)\Bigg/\left(\frac{\theta_2}{1-\theta_2}\right)\right\}^x \left(\frac{\theta_2}{1-\theta_2}\right)^{s} (1-\theta_1)^m (1-\theta_2)^n \\
&= {}_mC_x \cdot {}_nC_{s-x} \cdot \psi^x \omega^s \left(\frac{1}{1+\omega\psi}\right)^m \left(\frac{1}{1+\omega}\right)^n
\end{aligned}
$$

と (3.10) が示される．上式において変数は x であり，x を含まない項は定数であるので，全確率は 1 という条件から ${}_mC_x \cdot {}_nC_{s-x} \cdot \psi^x$ をその x に関する和 $\sum_k {}_mC_k \cdot {}_nC_{s-k} \cdot \psi^k$ で割ることにより (3.11) が導かれる．また，$\sum_k {}_mC_k \cdot {}_nC_{s-k} = {}_{m+n}C_s$ であるので (3.12) が成り立つ．

解答 3.2　$\theta_1 = \theta_2 = \theta$ のとき，$X/m - Y/n$ の標準偏差は

$$
SD\left[\frac{X}{m} - \frac{Y}{n}\right] = \sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\theta(1-\theta)}
$$

である．よって，歪度は

$$
\begin{aligned}
&\beta_1\left[\frac{X}{m} - \frac{Y}{n}\right] = E\left[\left\{\left(\frac{X}{m} - \frac{Y}{n}\right) - (\theta_1 - \theta_2)\right\}^3\right] \Bigg/ \left\{SD\left[\frac{X}{m} - \frac{Y}{n}\right]\right\}^3 \\
&= \left(\frac{1}{m^2} - \frac{1}{n^2}\right)\theta(1-\theta)(1-2\theta) \Bigg/ \left\{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\theta(1-\theta)}\right\}^3 \\
&= \left(\frac{n^2 - m^2}{m^2 n^2}\right)\theta(1-\theta)(1-2\theta) \Bigg/ \left\{\sqrt{\frac{m+n}{mn}\theta(1-\theta)}\right\}^3 \\
&= \frac{n-m}{\sqrt{mn(m+n)}} \cdot \frac{1-2\theta}{\sqrt{\theta(1-\theta)}}
\end{aligned}
$$

となる．尖度は

$$
\beta_2\left[\frac{X}{m} - \frac{Y}{n}\right] = E\left[\left\{\left(\frac{X}{m} - \frac{Y}{n}\right) - (\theta_1 - \theta_2)\right\}^4\right] \Bigg/ \left\{\left(\frac{1}{m} + \frac{1}{n}\right)\theta(1-\theta)\right\}^2 - 3
$$

$$=\left[3\left\{\left(\frac{1}{m}+\frac{1}{n}\right)\theta(1-\theta)\right\}^2+\left(\frac{1}{m^3}+\frac{1}{n^3}\right)\theta(1-\theta)\{1-6\theta(1-\theta)\}\right]\Bigg/\left\{\left(\frac{1}{m}+\frac{1}{n}\right)\theta(1-\theta)\right\}^2-3$$

$$=\frac{m^3+n^3}{mn(m+n)^2}\left\{\frac{1}{\theta(1-\theta)}-6\right\}$$

と求められる.

解答 3.3 帰無仮説の下での処置 1 の有効者数 X の分布は超幾何分布 $H(10, 13, 20)$ であり，$X = 3, . . . , 10$ の確率はそれぞれ右のようになる．$H_1 : \theta_1 > \theta_2$ に対する片側 P-値は

x	P(X=x)
3	0.0015
4	0.0271
5	0.1463
6	0.3251
7	0.3251
8	0.1463
9	0.0271
10	0.0015

$$P\text{-value} = \Pr(X \geq 8) = 0.1463 + 0.0271 + 0.0015 = 0.1749$$

となる．mid-P 値は

$$\text{mid-}P = 0.5 \times 0.1463 + 0.0271 + 0.0015 = 0.1018$$

である．$H_1 : \theta_1 \neq \theta_2$ に対する両側 P-値は，分布が左右対称であるので片側 P-値を 2 倍する流儀でも実現値の確率 $\Pr(X = 8) = 0.1463$ 以下の確率をすべて加える流儀でも同じ $0.1749 \times 2 = 0.3498$ となる.

解答 3.4 両群でのデータ数が等しく $m = n$ とすると $t = 2m - s$ であるので，公式 ${}_nC_k = {}_nC_{n-k}$ を用いると超幾何分布の確率は (3.15) より

$$\Pr(X=s-x)=\frac{{}_sC_{s-x}\times {}_{2m-s}C_{m-(s-x)}}{{}_NC_m}=\frac{{}_sC_x\times {}_{2m-s}C_{m-x}}{{}_NC_m}=\Pr(X=x)$$

となる.

解答 3.5 (3.22) については，$\hat{\theta}_1 = X/m$，$\hat{\theta}_2 = Y/n$ および $\hat{\theta}=(X+Y)/(m+n)$ より

$$\hat{\theta}_1-\hat{\theta}=\frac{X}{m}-\frac{X+Y}{m+n}=\frac{(m+n)X-m(X+Y)}{m(m+n)}=\frac{nX-mY}{m(m+n)}$$

$$\hat{\theta}_2-\hat{\theta}=\frac{Y}{n}-\frac{X+Y}{m+n}=\frac{(m+n)Y-n(X+Y)}{n(m+n)}=\frac{mY-nX}{n(m+n)}$$

$$\hat{\theta}_1-\hat{\theta}_2=\frac{X}{m}-\frac{Y}{n}=\frac{nX-mY}{mn}$$

となる．よって，

$$
\begin{aligned}
m(\hat{\theta}_1-\hat{\theta})^2+n(\hat{\theta}_2-\hat{\theta})^2 &= \frac{(nX-mY)^2}{m(m+n)^2}+\frac{(mY-nX)^2}{n(m+n)^2}\\
&= (mn)^2(\hat{\theta}_1-\hat{\theta}_2)^2\left\{\frac{1}{m(m+n)^2}+\frac{1}{n(m+n)^2}\right\}\\
&= \frac{(mn)^2(m+n)}{mn(m+n)^2}(\hat{\theta}_1-\hat{\theta}_2)^2=\frac{mn}{m+n}(\hat{\theta}_1-\hat{\theta}_2)^2=\frac{(\hat{\theta}_1-\hat{\theta}_2)^2}{\dfrac{1}{m}+\dfrac{1}{n}}
\end{aligned}
$$

より

$$
Z^2=\frac{(\hat{\theta}_1-\hat{\theta}_2)^2}{\left(\dfrac{1}{m}+\dfrac{1}{n}\right)\hat{\theta}(1-\hat{\theta})}=\frac{m(\hat{\theta}_1-\hat{\theta})^2+n(\hat{\theta}_2-\hat{\theta})^2}{\hat{\theta}(1-\hat{\theta})}
$$

となる．

(3.23) については，(3.22) の右辺の各項が

$$
\begin{aligned}
&\frac{m(\hat{\theta}_1-\hat{\theta})^2}{\hat{\theta}(1-\hat{\theta})}+\frac{n(\hat{\theta}_2-\hat{\theta})^2}{\hat{\theta}(1-\hat{\theta})}\\
&=\frac{m^2(\hat{\theta}_1-\hat{\theta})^2}{m\hat{\theta}}+\frac{m^2(\hat{\theta}_1-\hat{\theta})^2}{m(1-\hat{\theta})}+\frac{n^2(\hat{\theta}_2-\hat{\theta})^2}{n\hat{\theta}}+\frac{n^2(\hat{\theta}_2-\hat{\theta})^2}{n(1-\hat{\theta})}\\
&=\frac{(m\hat{\theta}_1-n_1\hat{\theta})^2}{m\hat{\theta}}+\frac{(m\hat{\theta}_1-m\hat{\theta})^2}{m(1-\hat{\theta})}+\frac{(n\hat{\theta}_2-n\hat{\theta})^2}{n\hat{\theta}}+\frac{(n\hat{\theta}_2-n\hat{\theta})^2}{n(1-\hat{\theta})}\\
&=\frac{(X-m\hat{\theta})^2}{m\hat{\theta}}+\frac{\{(m-X)-m(1-\hat{\theta})\}^2}{m(1-\hat{\theta})}+\frac{(Y-n\hat{\theta})^2}{n\hat{\theta}}+\frac{\{(n-Y)-n(1-\hat{\theta})\}^2}{n(1-\hat{\theta})}
\end{aligned}
$$

と変形されることから示される．

(3.24) については，(3.19) の Z の 2 乗を表 3.3 の記号で置き換えると，$N=m+n$ であるので

$$
Z^2=\frac{(\hat{\theta}_1-\hat{\theta}_2)^2}{\left(\dfrac{1}{m}+\dfrac{1}{n}\right)\hat{\theta}(1-\hat{\theta})}=\frac{\left(\dfrac{a}{m}-\dfrac{c}{n}\right)^2}{\dfrac{N}{mn}\dfrac{s}{N}\dfrac{t}{N}}=\frac{\dfrac{1}{(mn)^2}\{a(c+d)-c(a+b)\}^2}{\dfrac{st}{mnN}}=\frac{N(ad-bc)^2}{mnst}
$$

となる．

(3.25) については，超幾何分布 $H(m,\ s,\ N)$ の期待値と分散はそれぞれ $\dfrac{ms}{N}$，$\dfrac{mnst}{N^2(N-1)}$ であるので，超幾何分布をこれらの平均と分散を持つ正規分布で近似したときの連続修正を施した積分の端点は

$$\frac{(|a-ms/N|-0.5)^2}{mnst/\{N^2(N-1)\}}=\frac{(N-1)(|Na-ms|-N/2)^2}{mnst}$$
$$=\frac{(N-1)\{|(a+b+c+d)a-(a+b)(a+c)|-N/2\}^2}{mnst}=\frac{(N-1)(|ad-bc|-N/2)^2}{mnst}$$

となるので，$N-1\approx N$ として示される．

(3.26) に関しては，$\hat{\theta}_1=a/m$，$\hat{\theta}_2=c/n$ および $\hat{\theta}=(a+c)/(m+n)$ を用いると $(z')^2$ が (3.25) に一致することが次のように示される．

$$(z')^2=\frac{\left\{\left|\frac{a}{m}-\frac{c}{n}\right|-\frac{1}{2}\left(\frac{1}{m}+\frac{1}{n}\right)\right\}^2}{\left(\frac{1}{m}+\frac{1}{n}\right)\frac{a+c}{m+n}\left(1-\frac{a+c}{m+n}\right)}=\frac{\frac{1}{(mn)^2}\left\{|na-mc|-\frac{1}{2}(m+n)\right\}^2}{\frac{m+n}{mn}\cdot\frac{a+c}{m+n}\cdot\frac{(m+n)-(a+c)}{m+n}}=\frac{N(|ad-bc|-N/2)^2}{mnst}$$

解答 3.6　各処置での有効率はそれぞれ $\hat{\theta}_1=8/10=0.8$，$\hat{\theta}_2=5/10=0.5$ であり，H_0 : $\theta_1=\theta_2\ (=\theta)$ の下での θ の推定値は $\hat{\theta}=13/20=0.65$ である．よって

$$z^*=\frac{0.8-0.5}{\sqrt{\left(\frac{1}{10}+\frac{1}{10}\right)0.65\times 0.35}}\approx 1.4064$$

となる．片側 P-値は $P_1=\Pr(Z\geq 1.4064)\approx 0.0798$ となる．$w^*=(z^*)^2=1.978$ であり，両側 P-値は $P_2=\Pr(W\geq 1.978)\approx 0.1596$ となる．イェーツの補正を施した w' は

$$w'=\frac{20(|8\times 5-2\times 5|-20/2)^2}{10\times 10\times 13\times 7}=\frac{20\times(30-10)^2}{10\times 10\times 13\times 7}=0.8791$$

となり，P-値は $P_{Yates}=\Pr(W\geq 0.8791)\approx 0.3484$ となる．P_{Yates} はフィッシャー検定の両側 P-値 0.3498 に極めて近い．

解答 3.7　$m=n=10$，$x=8$，$y=5$ であるので，$d=8/10-5/10=0.8-0.5=0.3$ であり，標準誤差は

$$SE[d]=\sqrt{0.8(1-0.8)/10+0.5(1-0.5)/10}=\sqrt{0.041}\approx 0.2025$$

となる．また，$\tilde{d}_1=\frac{8+1}{10+2}-\frac{5+1}{10+2}=\frac{9}{12}-\frac{6}{12}=0.75-0.5=0.25$ であり，形式的に計算した標準誤差は

$$SE[\tilde{d}_1]=\sqrt{0.75(1-0.75)/12+0.5(1-0.5)/12}\approx\sqrt{0.0365}\approx 0.1909$$

となる．

解答 3.8

$$\frac{mn}{m+n}SE_{Wald}{}^2=\frac{n}{m+n}\hat{\theta}_1(1-\hat{\theta}_1)+\frac{m}{m+n}\hat{\theta}_2(1-\hat{\theta}_2)=(1-\gamma)\hat{\theta}_1(1-\hat{\theta}_1)+\gamma\hat{\theta}_2(1-\hat{\theta}_2)$$

であり， $\hat{\theta}=\gamma\hat{\theta}_1+(1-\gamma)\hat{\theta}_2$ より

$$\begin{aligned}&\frac{mn}{m+n}SE_{Score}{}^2=\{\gamma\hat{\theta}_1+(1-\gamma)\hat{\theta}_2\}[1-\{\gamma\hat{\theta}_1+(1-\gamma)\hat{\theta}_2\}]\\&=\{\hat{\theta}_2+\gamma(\hat{\theta}_1-\hat{\theta}_2)\}\{(1-\hat{\theta}_2)-\gamma(\hat{\theta}_1-\hat{\theta}_2)\}\\&=\hat{\theta}_2(1-\hat{\theta}_2)+(1-2\hat{\theta}_2)\gamma(\hat{\theta}_1-\hat{\theta}_2)-\gamma^2(\hat{\theta}_1-\hat{\theta}_2)^2\end{aligned}$$

となる．ここで

$$\begin{aligned}&(1-2\hat{\theta}_2)\gamma(\hat{\theta}_1-\hat{\theta}_2)=(1-\hat{\theta}_1-\hat{\theta}_2+\hat{\theta}_1-\hat{\theta}_2)\gamma(\hat{\theta}_1-\hat{\theta}_2)\\&=\gamma(\hat{\theta}_1-\hat{\theta}_2)(1-\hat{\theta}_1-\hat{\theta}_2)+\gamma(\hat{\theta}_1-\hat{\theta}_2)^2=\gamma\{\hat{\theta}_1(1-\hat{\theta}_1)-\hat{\theta}_2(1-\hat{\theta}_2)\}+\gamma(\hat{\theta}_1-\hat{\theta}_2)^2\end{aligned}$$

に注意すると，

$$\frac{mn}{m+n}SE_{Score}{}^2=(1-\gamma)\hat{\theta}_2(1-\hat{\theta}_2)+\gamma\hat{\theta}_1(1-\hat{\theta}_1)+\gamma(1-\gamma)(\hat{\theta}_1-\hat{\theta}_2)^2$$

となる．よって，

$$\begin{aligned}&\frac{mn}{m+n}(SE_{Score}{}^2-SE_{Wald}{}^2)\\&=(1-\gamma)\hat{\theta}_2(1-\hat{\theta}_2)+\gamma\hat{\theta}_1(1-\hat{\theta}_1)+\gamma(1-\gamma)(\hat{\theta}_1-\hat{\theta}_2)^2-\{(1-\gamma)\hat{\theta}_1(1-\hat{\theta}_1)+\gamma\hat{\theta}_2(1-\hat{\theta}_2)\}\\&=\gamma(1-\gamma)(\hat{\theta}_1-\hat{\theta}_2)^2+(1-2\gamma)\{\hat{\theta}_2(1-\hat{\theta}_2)-\hat{\theta}_1(1-\hat{\theta}_1)\}\end{aligned}$$

を得る．

解答 3.9　$SE_{Wald}>SE_{Score}$ となるのは，(3.43) より γ が小さくて $\hat{\theta}_2(1-\hat{\theta}_2)<\hat{\theta}_1(1-\hat{\theta}_1)$ の場合である．具体的に $m=6$, $n=50$ で $x=3$, $y=5$ では $\hat{\theta}_1=3/6=0.5$, $\hat{\theta}_2=5/50=0.1$ および $\hat{\theta}=8/56=1/7=0.143$ であり，$SE_{Score}-SE_{Wald}=-0.0573$ と負になる．

解答 3.10

$m=10$，$n=10$，$x=8$，$y=4$ であるので $d=0.3$ であり，

$$SE_{Wald}[d]=\sqrt{0.8(1-0.8)/10+0.5(1-0.5)/10}=\sqrt{0.0410}\approx 0.2025$$

より，ワルド型の 95%信頼区間は

$$0.3\pm 1.96\times 0.2025=(-0.0969,\,0.6969)$$

となる．スコア型の信頼区間は，$\hat{\theta}=(8+5)/(10+10)=0.65$ より

$$SE_{Score}[d] = \sqrt{\left(\frac{1}{10}+\frac{1}{10}\right)0.65(1-0.65)} \approx \sqrt{0.0455} \approx 0.2133$$

であるので

$$0.3 \pm 1.96 \times 0.2133 = (-0.1181, 0.7181)$$

と求められる.

処置 1 および処置 2 でのスコア型の 95%信頼区間が 2.3.4 節の (2.28) よりそれぞれ (0.4902, 0.9433), (0.2366, 0.7634) となる. よって, (3.45) および (3.46) の標準誤差部分はそれぞれ

$$\sqrt{\frac{0.4902(1-0.4902)}{10}+\frac{0.7634(1-0.7634)}{10}} \approx 0.2075,$$

$$\sqrt{\frac{0.9433(1-0.9433)}{10}+\frac{0.2366(1-0.2366)}{8}} \approx 0.1530$$

であるので, ハイブリッドスコア型 95%信頼区間の上下限は (3.45) および (3.46) より

$$\delta_L = 0.3 - 1.96 \times 0.2075 \approx -0.1067, \quad \delta_U = 0.3 + 1.96 \times 0.1530 \approx 0.5999$$

と求められる.

また, 仮想的に各度数に 1 ずつを加える方法では $\tilde{d}_1 = 0.25$ であり, 標準誤差は $SE[\tilde{d}_1] \approx 0.1909$ であるので, 95%信頼区間は (3.47) より

$$\tilde{d}_1 \pm z(\alpha/2)SE[\tilde{d}_1] = 0.25 \pm 1.96 \times 0.1909 = (-0.1242, 0.6242)$$

となる.

解答 3.11 n 回の試行で生起確率 θ の事象を 1 回以上観測する確率は $1-(1-\theta)^n$ で与えられる. よって, $1-(1-\theta)^n \geq 0.95$ より $(1-\theta)^n \leq 0.05$ であるので, 両辺の対数を取って $n\log(1-\theta) \leq \log 0.05$ を得る. θ が小さいとき $\log(1-\theta) \approx -\theta$ であるので $\log 0.05 \approx -3$ となり, 近似的に $-n\theta \leq -3$ が得られ, $n \geq 3/\theta$ となる.

解答 3.12 ベイズの定理 $\Pr(A \mid B) = \Pr(B \mid A)\Pr(A)/\Pr(B)$ より

$$\frac{\eta_1/(1-\eta_1)}{\eta_2/(1-\eta_2)} = \frac{\Pr(A \mid B)/\Pr(A^c \mid B)}{\Pr(A \mid B^c)/\Pr(A^c \mid B^c)}$$

$$=\frac{\{\Pr(B\mid A)\Pr(A)/\Pr(B)\}/\{\Pr(B\mid A^c)\Pr(A^c)/\Pr(B)\}}{\{\Pr(B^c\mid A)\Pr(A)/\Pr(B^c)\}/\{\Pr(B^c\mid A^c)\Pr(A^c)/\Pr(B^c)\}}$$

$$=\frac{\Pr(B\mid A)/\Pr(B\mid A^c)}{\Pr(B^c\mid A)/\Pr(B^c\mid A^c)}=\frac{\Pr(B\mid A)/\Pr(B^c\mid A)}{\Pr(B\mid A^c)/\Pr(B^c\mid A^c)}=\psi$$

となる.

解答 3.13　オッズ比は ψ = (0.003/0.997)/(0.001/0.999) ≈ 3.006018 と求められる. リスク比は θ_1/θ_2 = 0.003/0.001 = 3 であり, 両者はきわめて近い. ちなみに近似の次数を上げて $\theta/(1-\theta) \approx \theta + \theta^2$ とすると $0.003/0.997 \approx 0.003 + (0.003)^2 = 0.003009$ および $0.001/0.999 \approx 0.001 + (0.001)^2 = 0.001001$ であるので 0.003009 /0.001001 ≈ 3.005994 となる.

解答 3.14　標本オッズ比は $y = (2 \times 4)/(8 \times 4) = 1/4 = 0.25$ であり, 対数オッズ比は $\log 0.25 \approx -1.386$ となる. 標準誤差は

$$SE[\log y]=\sqrt{\frac{1}{2}+\frac{1}{8}+\frac{1}{4}+\frac{1}{4}}=\sqrt{\frac{9}{8}}\approx 1.061$$

であるので, 対数オッズ比 $\log\,\psi$ の近似的な 95%信頼区間は $-1.386 \pm 1.96 \times 1.061 \approx (-3.465, 0.693)$ となる. よってオッズ比 ψ の 95%信頼区間は (exp[−3.465], exp[0.693]) = (0.031, 1.999) となる. 正確な 95%信頼区間は数値計算により (0.017, 2.874) となる.

解答 3.15　$\psi = 1$ のとき非心超幾何分布 (3.56) は通常の超幾何分布 $H(m, s, N)$ となり, (A, B) セル度数 X の期待値は $a_0 = ms/N$ となる（2.2.4 節参照）. また, その他のセル度数の期待値はそれぞれ $b_0 = m - a_0 = mt/N$, $c_0 = s - a_0 = ns/N$, $d_0 = s - n - a_0 = nt/N$ となる. よって

$$SE_0=\sqrt{\frac{N}{ms}+\frac{N}{mt}+\frac{N}{ns}+\frac{N}{nt}}=\sqrt{\frac{N(nt+ns+mt+ms)}{mnst}}=\sqrt{\frac{N^3}{mnst}}=N\sqrt{\frac{N}{mnst}}$$

を得る.

解答 3.16　各確率の推定値は $\hat{\theta}_j = x_j/n_j$ および $\xi_j = n_j\bar{\theta}$ であるので, (3.61) 式の分母分子を n_j で割って整理すると

$$
\begin{aligned}
y &= \sum_{j=1}^{m}\frac{(x_j-\xi_j)^2}{\xi_j}+\sum_{j=1}^{m}\frac{\{(n_j-x_j)-(n_j-\xi_j)\}^2}{n_j-\xi_j} \\
&= \sum_{j=1}^{m}\frac{n_j(x_j/n_j-\xi_j/n_j)^2}{\xi_j/n_j}+\sum_{j=1}^{m}\frac{n_j(x_j/n_j-\xi_j/n_j)^2}{1-\xi_j/n_j} \\
&= \sum_{j=1}^{m}n_j(\hat{\theta}_j-\overline{\theta})^2\left\{\frac{1}{\overline{\theta}}+\frac{1}{1-\overline{\theta}}\right\}=\frac{1}{\overline{\theta}(1-\overline{\theta})}\sum_{j=1}^{m}n_j(\hat{\theta}_j-\overline{\theta})^2
\end{aligned}
$$

となる．

解答 3.17　すべての統計量の分母は同じ $\overline{\theta}(1-\overline{\theta})$ であるので，分子のみ計算する．まず

$$
\overline{\theta}=\frac{1}{N}\sum_{j=1}^{m}x_j=\frac{1}{N}\left(\sum_{j=1}^{k}x_j+\sum_{j=k+1}^{m}x_j\right)=\frac{1}{N}(n^{(1)}\overline{\theta}^{(1)}+n^{(2)}\overline{\theta}^{(2)})
$$

であり，

$$
\overline{\theta}^{(1)}=\frac{1}{n^{(1)}}\sum_{j=1}^{k}x_j=\frac{1}{n^{(1)}}\sum_{j=1}^{k}n_j\hat{\theta}_j,\quad \overline{\theta}^{(2)}=\frac{1}{n^{(2)}}\sum_{j=k+1}^{m}x_j=\frac{1}{n^{(2)}}\sum_{j=k+1}^{m}n_j\hat{\theta}_j
$$

であることに注意する．これらより

$$
\begin{aligned}
&\sum_{j=1}^{m}n_j(\hat{\theta}_j-\overline{\theta})^2=\sum_{j=1}^{k}n_j\{(\hat{\theta}_j-\overline{\theta}^{(1)})+(\overline{\theta}^{(1)}-\overline{\theta})\}^2+\sum_{j=k+1}^{m}n_j\{(\hat{\theta}_j-\overline{\theta}^{(2)})+(\overline{\theta}^{(2)}-\overline{\theta})\}^2 \\
&=\sum_{j=1}^{k}n_j(\hat{\theta}_j-\overline{\theta}^{(1)})^2+2(\overline{\theta}^{(1)}-\overline{\theta})\sum_{j=1}^{k}n_j(\hat{\theta}_j-\overline{\theta}^{(1)})+(\overline{\theta}^{(1)}-\overline{\theta})^2\sum_{j=1}^{k}n_j \\
&\quad+\sum_{j=k+1}^{m}n_j(\hat{\theta}_j-\overline{\theta}^{(2)})^2+2(\overline{\theta}^{(2)}-\overline{\theta})\sum_{j=1}^{k}n_j(\hat{\theta}_j-\overline{\theta}^{(2)})+(\overline{\theta}^{(2)}-\overline{\theta})^2\sum_{j=1}^{k}n_j \\
&=\sum_{j=1}^{k}n_j(\hat{\theta}_j-\overline{\theta}^{(1)})^2+0+n^{(1)}\left\{\overline{\theta}^{(1)}-\frac{1}{N}(n^{(1)}\overline{\theta}^{(1)}+n^{(2)}\overline{\theta}^{(2)})\right\}^2 \\
&\quad+\sum_{j=k+1}^{m}n_j(\hat{\theta}_j-\overline{\theta}^{(2)})^2+0+n^{(2)}\left\{\overline{\theta}^{(2)}-\frac{1}{N}(n^{(1)}\overline{\theta}^{(1)}+n^{(2)}\overline{\theta}^{(2)})\right\}^2 \\
&=\sum_{j=1}^{k}n_j(\hat{\theta}_j-\overline{\theta}^{(1)})^2+\frac{n^{(1)}\{n^{(2)}\}^2}{N^2}(\overline{\theta}^{(1)}-\overline{\theta}^{(2)})^2+\sum_{j=k+1}^{m}n_j(\hat{\theta}_j-\overline{\theta}^{(2)})^2+\frac{\{n^{(1)}\}^2n^{(2)}}{N^2}(\overline{\theta}^{(1)}-\overline{\theta}^{(2)})^2 \\
&=y^{(1)}+y^{(2)}+\frac{n^{(1)}n^{(2)}\{n^{(1)}+n^{(2)}\}}{N^2}(\overline{\theta}^{(1)}-\overline{\theta}^{(2)})^2=y^{(1)}+y^{(2)}+y_{diff}
\end{aligned}
$$

を得る．

解答 3.18　すべての統計量の分母は同じ $\overline{\theta}(1-\overline{\theta})$ であるので，それ以外の部分のみ計算する．予測値は $\theta_j{*}=\overline{\theta}+b(h_j-\overline{h})$ であるので，カイ 2 乗統計量 y の定義

(3.62) より

$$\sum_{j=1}^{m} n_j(\hat{\theta}_j-\overline{\theta})^2=\sum_{j=1}^{m} n_j\{(\hat{\theta}_j-\theta_j^*)-(\theta_j^*-\overline{\theta})\}^2$$
$$=\sum_{j=1}^{m} n_j(\hat{\theta}_j-\theta_j^*)^2-2\sum_{j=1}^{m} n_j(\hat{\theta}_j-\theta_j^*)(\theta_j^*-\overline{\theta})+\sum_{j=1}^{m} n_j(\theta_j^*-\overline{\theta})^2$$
$$=\sum_{j=1}^{m} n_j(\hat{\theta}_j-\theta_j^*)^2-2\sum_{j=1}^{m} n_j\{\hat{\theta}_j-\overline{\theta}-b(h_j-\overline{h})\}\{b(h_j-\overline{h})\}+\sum_{j=1}^{m} n_j\{b(h_j-\overline{h})\}^2$$

であるが，最後の式の第 2 項の和は $\sum_{j=1}^{m} n_j(\hat{\theta}_j-\overline{\theta})(h_j-\overline{h})-b\sum_{j=1}^{m} n_j(h_j-\overline{h})^2$ であるが，b の定義よりこれは 0 になる．よって予式が成り立つ．

解答 3.19　条件付き確率の差は

$$\xi_1-\xi_2=\frac{\pi_{12}-\pi_{21}}{\pi_{12}+\pi_{21}}=\frac{\pi_{12}/\pi_{21}-1}{\pi_{12}/\pi_{21}+1}=\frac{\psi-1}{\psi+1}$$

とオッズ比 ψ の単調関数として表わすことができる．よって，$\xi_1-\xi_2$ が等しければオッズ比 ψ も等しくなる．逆も同様である．

解答 3.20　$X \sim B(b+c, 1/2)$ とすると，近似的に $X \sim N((b+c)/2, (b+c)/4)$ である．$b \geq (b+c)/2$ すなわち $b \geq c$ とすると，連続修正を施した場合の確率計算は

$$\Pr(X \geq b-0.5)=\Pr\left(\frac{X-(b+c)/2}{\sqrt{(b+c)/4}} \geq \frac{b-0.5-(b+c)/2}{\sqrt{(b+c)/4}}\right)=\Pr\left(Z \geq \frac{(b-c)-1}{\sqrt{b+c}}\right)$$

より，端点を 2 乗すると $\{(b-c)-1\}^2/(b+c)$ となる．逆に $b \leq (b+c)/2$ すなわち $c \geq b$ とすると，連続修正を施した場合の確率計算は

$$\Pr(X \leq b+0.5)=\Pr\left(\frac{X-(b+c)/2}{\sqrt{(b+c)/4}} \leq \frac{b+0.5-(b+c)/2}{\sqrt{(b+c)/4}}\right)=\Pr\left(Z \leq -\frac{(c-b)-1}{\sqrt{b+c}}\right)$$

より端点を 2 乗して $\{(c-b)-1\}^2/(b+c)$ となる．これらより結局 $(|b-c|-1)^2/(b+c)$ を得る．

解答 3.21　マクネマー検定で $100\alpha\%$有意になることと信頼区間 $(b-c)/\ \pm$ $z(\alpha/2)\sqrt{b+c}/n$ が 0 を含まないことが同値であることを示す．信頼区間の下限については $(b-c)/n-z(\alpha/2)\sqrt{b+c}/n > 0$ より $(b-c)/n>z(\alpha/2)\sqrt{b+c}/n$ であるのでこの両辺を n 倍して 2 乗し，$(b-c)^2/(b+c) > (z(\alpha/2))^2$ を得る．これは検定で $100\alpha\%$有意であることを示す．上限についても同様の結果が得られる．

解答 3.22　不等式 $\dfrac{(b-c-n\delta_0)^2}{(b+c)-n{\delta_0}^2}<\chi^2(\alpha)$ を変形した

$$(b-c-n\delta_0)^2<\chi^2(\alpha)\{(b+c)-n\delta_0{}^2\}$$

より

$$n\{n+\chi^2(\alpha)\}\delta_0{}^2-2(b-c)n\delta_0+(b-c)^2-\chi^2(\alpha)(b+c)<0$$

と δ_0 の 2 次式となるので，解の公式より区間の端点は以下のようになる：

$$\frac{(b-c)}{n+\chi^2(\alpha)}\pm\frac{\sqrt{(b-c)^2-\{n+\chi^2(\alpha)\}\{(b-c)^2-\chi^2(\alpha)(b+c)\}/n}}{n+\chi^2(\alpha)}$$

$$\frac{(b-c)}{n+\chi^2(\alpha)}\pm\frac{\sqrt{-\chi^2(\alpha)(b-c)^2/n+\{n+\chi^2(\alpha)\}\chi^2(\alpha)(b+c)/n}}{n+\chi^2(\alpha)}$$

$$\frac{(b-c)}{n+\chi^2(\alpha)}\pm z(\alpha/2)\frac{\sqrt{[\{n+\chi^2(\alpha)\}(b+c)-(b-c)^2]/n}}{n+\chi^2(\alpha)}$$

解答 3.23　(3.101) の公式 $n^*=\dfrac{n}{4}\left\{1+\sqrt{1+\dfrac{4}{n|\theta_1-\theta_2|}}\right\}^2$ より n が大きいとき

$$n^*=\frac{n}{4}\left\{2+\frac{4}{n|\theta_1-\theta_2|}+2\sqrt{1+\frac{4}{n|\theta_1-\theta_2|}}\right\}$$

$$\approx\frac{n}{4}\left\{2+\frac{4}{n|\theta_1-\theta_2|}+2\left(1+\frac{2}{n|\theta_1-\theta_2|}\right)\right\}=n+\frac{2}{|\theta_1-\theta_2|}$$

となる．同様に (3.104) より n_1 が大きいとき

$$n_1{}^*=\frac{n_1}{4}\left\{1+\sqrt{1+\frac{2(1+r)}{n_1r|\theta_1-\theta_2|}}\right\}^2=\frac{n_1}{4}\left\{2+\frac{2(1+r)}{n_1r|\theta_1-\theta_2|}+\sqrt{1+\frac{2(1+r)}{n_1r|\theta_1-\theta_2|}}\right\}$$

$$\approx\frac{n_1}{4}\left\{2+\frac{2(1+r)}{n_1r|\theta_1-\theta_2|}+2\left(1+\frac{2(1+r)}{n_1r|\theta_1-\theta_2|}\right)\right\}=n_1+\frac{1+r}{r|\theta_1-\theta_2|}$$

を得る．

解答 3.24　(3.107) より

$$z(\alpha/2)\sqrt{n(\pi_{12}+\pi_{21})}-n(\pi_{12}-\pi_{21})=z(1-\beta)\sqrt{n\{(\pi_{12}+\pi_{21})-(\pi_{12}-\pi_{21})^2\}}$$

であるので，移項して $\sqrt{n}$ で割り $z(b)=-z(1-\beta)$ に注意すると

$$\sqrt{n}=\frac{z(\alpha/2)\sqrt{\pi_{12}+\pi_{21}}+z(\beta)\sqrt{(\pi_{12}+\pi_{21})-(\pi_{12}-\pi_{21})^2}}{\pi_{12}-\pi_{21}}$$

となり，2 乗して (3.108) を得る．