Notes on the odds ratio and the δ-method

Let $X \sim B(n, p)$ and $Y \sim B(m, q)$, independent; $0 < p,\ q < 1$, while n and m are large. The true odds ratio is

$$\phi = \frac{p/(1-p)}{q/(1-q)}.$$

This parameter is estimated by the statistic

$$\hat{\phi} = \frac{\hat{p}/(1-\hat{p})}{\hat{q}/(1-\hat{q})},$$

where

$$\hat{p} = X/n = p + \delta_n = p + \sqrt{p(1-p)}\xi_n/\sqrt{n} \tag{1}$$

and

$$\hat{q} = Y/m = q + \epsilon_m = q + \sqrt{q(1-q)}\zeta_m/\sqrt{m} \tag{2}$$

with $\xi_n,\ \zeta_m$ independent and asymptotically $N(0, 1)$. Equation (1) defines δ_n and ξ_n in terms of X and n; likewise, (2) defines ϵ_m and ζ_m in terms of Y and m.

Asymptotically, $\log \hat{\phi}$ is normal, with asymptotic mean $\log \phi$ and asymptotic variance

$$\frac{1}{np} + \frac{1}{n(1-p)} + \frac{1}{mq} + \frac{1}{m(1-q)}. \tag{3}$$

This can be argued by the so-called δ-method, since

$$\begin{aligned}
\log \hat{\phi} - \log \hat{\phi} &= \log \frac{\hat{p}}{p} - \log \frac{1-\hat{p}}{1-p} - \log \frac{\hat{q}}{q} + \log \frac{1-\hat{q}}{1-q} \\
&= \log \Big(1 + \frac{\delta_n}{p}\Big) - \log \Big(1 - \frac{\delta_n}{1-p}\Big) - \log \Big(1 + \frac{\epsilon_m}{q}\Big) + \log \Big(1 - \frac{\epsilon_m}{1-q}\Big) \\
&\doteq \Big(\frac{1}{p} + \frac{1}{1-p}\Big)\delta_n - \Big(\frac{1}{q} + \frac{1}{1-q}\Big)\epsilon_m \\
&= \frac{1}{\sqrt{p(1-p)}}\frac{\xi_n}{\sqrt{n}} - \frac{1}{\sqrt{q(1-q)}}\frac{\zeta_m}{\sqrt{m}}.
\end{aligned}$$

The variance of the last displayed expression is (3), which is usually estimated as

$$\frac{1}{X} + \frac{1}{n-X} + \frac{1}{Y} + \frac{1}{m-Y}.$$

In finite samples, $\hat{\phi}$ has infinite mean, since $Y = 0$ has positive probability.