Remarks on the Example: Estimation of Prevalence Using an Imperfect Test

The remarks that follow pertain to the first part of the example on page 393 that illustrates a computation of the frequentist confidence interval for the prevalence of disease in the population, conditional on the sensitivity and specificity of the test.

The notation used in the book and in this file is as follows:

$\Pr(Y=1)$: the proportion of subjects in the population that have the disease. This is an unobserved quantity and the focus of inference.

$\Pr\left(\widehat{Y}=1\right)$: the probability of a positive outcome (individual is declared diseased) based on an imperfect test.

$\widehat{\Pr}\left(\widehat{Y}=1\right)$: the estimator of $\Pr\left(\widehat{Y}=1\right)$.

$$E\left(\widehat{\Pr}\left(\widehat{Y}=1\right)\right)=\Pr\left(\widehat{Y}=1\right)=(1-spe)+(sen+spe-1)\Pr(Y=1) \tag{1}$$

as indicated below eqn. (9.29) on page 392, where specificity and sensitivity of the test are respectively

$spe : \Pr\left(\widehat{Y}=0|Y=0\right)$,

$sen : \Pr\left(\widehat{Y}=1|Y=1\right)$.

The sampling distribution of T, number of positive outcomes out of n individuals tested is binomially distributed $Bi\left(n,\Pr\left(\widehat{Y}=1\right)\right)$. The example uses this result and computes initial bounds for the confidence interval for $\Pr\left(\widehat{Y}=1\right)$, which are then adjusted based on eqn. (1) to retrieve a confidence interval for $\Pr(Y=1)$. The initial bounds are approximated using the estimate $\widehat{\Pr}\left(\widehat{Y}=1\right)=T/n$ in lieu of the true parameter $\Pr\left(\widehat{Y}=1\right)$. As indicated on page 393, the unadjusted lower and upper bounds are obtained using the quantiles of the binomial distribution:

```
lb <- qbinom(0.025,n,(T/n))
ub <- qbinom(0.975,n,(T/n))
```

The adjusted bounds that follow give rise to the approximate confidence interval displayed in eqn. (9.31) on page 393.

This note describes two alternative approaches that can be followed to arrive at frequentist confidence intervals for $\Pr(Y=1)$. One is well known and is based on the asymptotic normal approximation; the other is perhaps less well known and is based on the Clopper-Pearson exact confidence interval for the binomial proportion (Clopper and Pearson, 1934). Both approaches are illustrated using the same example discussed in the book on page 393.

Normal Approximation

The starting point is the sampling distribution of T,

$$T|n, \Pr\left(\widehat{Y}=1\right) \sim Bi\left(n, \Pr\left(\widehat{Y}=1\right)\right). \tag{2}$$

In order to use more compact notation it will be convenient to define

$$\begin{aligned}
\Pr\left(\widehat{Y}=1\right) &= \phi, \\
\Pr(Y=1) &= \theta, \\
(1-spe) &= a, \\
(sen+spe-1) &= b,
\end{aligned}$$

and therefore using eqn. (1), $\phi = a + b\,\theta$. Given the model,

$$\begin{aligned}
E(T) &= n\phi, \\
Var(T) &= n\phi(1-\phi), \\
\widehat{\Pr}\left(\widehat{Y}=1\right) &= \widehat{\phi} = \frac{T}{n}, \\
E\left(\widehat{\phi}\right) &= \phi, \\
Var\left(\widehat{\phi}\right) &= \frac{\phi(1-\phi)}{n}, \\
\widehat{Var}\left(\widehat{\phi}\right) &= \frac{\widehat{\phi}\left(1-\widehat{\phi}\right)}{n},
\end{aligned}$$

and using the normal approximation,

$$\frac{\widehat{\phi}-\phi}{se\left(\widehat{\phi}\right)} \sim N(0,1) \tag{3}$$

where

$$se\left(\widehat{\phi}\right) = \sqrt{\widehat{Var}\left(\widehat{\phi}\right)}.$$

The example in the book assumes that the imperfect test is performed on $n = 10,000$ subjects and that $T = 900$ show a positive result. The sensitivity and specificity of the test are known to be 085 and 0.95, respectively. Then,

$$
\begin{aligned}
a &= 1 - spe = 1 - 0.85 = 0.05, \\
b &= spe + sen - 1 = 0.95 + 0.85 - 1 = 0.8, \\
n &= 10,000, \\
T &= 900, \\
\widehat{\phi} &= 0.09, \\
\widehat{Var}\left(\widehat{\phi}\right) &= \frac{0.09 \times 0.91}{10,000}, \\
se\left(\widehat{\phi}\right) &= \left[\frac{0.09 \times 0.91}{10,000}\right]^{\frac{1}{2}}, \\
\widehat{\theta} &= \frac{\widehat{\phi} - a}{b} = 0.05.
\end{aligned}
$$

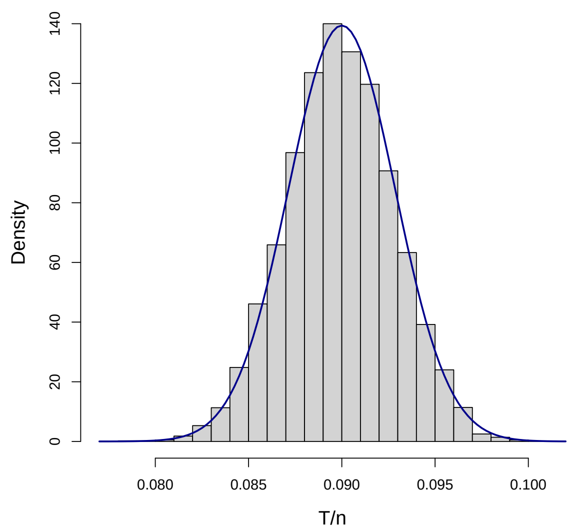

Figure 1: Histogram of the distribution of repeated draws T/n from a binomial distribution $Bi(n = 10,000, \Pr(\hat{Y} = 1) = 0.09)$ with a normal distribution $N(\mu = 0.09, \sigma^2 = (0.09 \times 0.91/10,000))$ overlaid.

Confidence interval for $\phi = a + b\theta$, given (a, b), follows from (3) based on the normal approximation and eventually for θ:

$$
\begin{aligned}
\Pr\left(-1.96 < \frac{\widehat{\phi}-\phi}{se\left(\widehat{\phi}\right)} < 1.96\right) &= 0.95, && \text{(4a)}\\
\Pr\left(\widehat{\phi} - 1.96\, se\left(\widehat{\phi}\right) < \phi < \widehat{\phi} + 1.96\, se\left(\widehat{\phi}\right)\right) &= 0.95, && \text{(4b)}\\
\Pr\left(\widehat{\phi} - 1.96\, se\left(\widehat{\phi}\right) < a + b\theta < \widehat{\phi} + 1.96\, se\left(\widehat{\phi}\right)\right) &= 0.95, && \text{(4c)}\\
\Pr\left(\frac{\widehat{\phi} - 1.96\, se\left(\widehat{\phi}\right) - a}{b} < \theta < \frac{\widehat{\phi} + 1.96\, se\left(\widehat{\phi}\right) - a}{b}\right) &= 0.95, && \text{(4d)}\\
\Pr(0.043 < \theta < 0.057) &= 0.95, && \text{(4e)}
\end{aligned}
$$

the same result reported in eqn. (9.31) on page 393 of the book.

The approximation based on (4b) is not guaranteed to generate bounds within the allowed parameter space $\phi \in (0, 1)$ particularly when ϕ takes values close to 0 or 1. This is aggravated further in the case of the confidence interval for θ given by (4d). If (lb, ub) is the confidence interval for ϕ, a common heuristic is to truncate the confidence interval for θ as

$$\max\left[0, \frac{lb-a}{b}\right], \min\left[1, \frac{ub-a}{b}\right] \tag{5}$$

(Diggle, 2011). Figure 1 displays the histogram of draws T/n, $T = 900, n = 10,000$ from $T \sim Bi(n, \phi = 0.09)$ and a normal distribution $N(\mu = 0.09, \sigma^2 = (0.09 \times 0.91/10,000)$ overlaid. With this size of sample n the normal approximation gives adequate results despite the extreme value of ϕ.

The Clopper-Pearson exact confidence interval

The Clopper-Pearson interval for the proportion ϕ in the population contains the values of ϕ that are not rejected by the test at level α ($\alpha = 0.05$ for the 95% confidence interval). Assume that for a given binomial experiment, x successes are observed out of n trials. Then the Clopper-Pearson lower bound for a two-sided test retrieves the smallest possible value of ϕ, ϕ_{lb}, for having an $\alpha/2$ probability of observing at least x successes. Given x, the lower bound is given by the value of ϕ_{lb} satisfying

$$\sum_{k=x}^{n} \binom{n}{k} \phi_{lb}^{k}(1-\phi_{lb})^{n-k} \geq \alpha/2. \tag{6}$$

The upper bound provides the largest possible value of ϕ, ϕ_{ub}, for having an $\alpha/2$ probability of observing no more than x successes. Given x, the upper bound is given by the value of ϕ_{ub} satisfying

$$\sum_{k=0}^{x} \binom{n}{k} \phi_{ub}^{k}(1-\phi_{ub})^{n-k} \geq \alpha/2. \tag{7}$$

Numerical methods are needed to obtain ϕ_{lb} and ϕ_{ub}. The computation is greatly simplified using the equality from Johnson et al., 2005:

$$\sum_{k=x}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \int_0^1 f(u, x, n-x+1)du \tag{8}$$

where f is the density function of the beta distribution $Be(a, b)$. When (8) is used in (6) and (7), the Clopper-Pearson bounds for the interval are given by the quantiles of the beta distributions:

$$(\phi_{lb}, \phi_{ub}) = (Be(\alpha/2, x, n-x+1), Be(1-\alpha/2, x+1, n-x)). \tag{9}$$

As illustrated below, these bounds are obtained using `R` for example, with one line of code. Closed form solutions are obtained when x is either 0 or 1. When $x = 0$ the interval is $\left(0, 1-(\alpha/2)^{\frac{1}{n}}\right)$ and when $x = n$, the interval is $\left((\alpha/2)^{\frac{1}{n}}, 1\right)$.

Returning to the example in the book with $n = 10,000$ and $T = 900$, the Clopper-Pearson interval for ϕ is obtained from

```
lb <- qbeta(0.025,900,9101)
ub <- qbeta(0.975,901,9100)
```

This results in an interval for a positive test given by

$$\Pr(0.08446093 < \phi < 0.09577916) = 0.95$$

and using (5), the interval for the prevalence in the population is

$$\begin{aligned}\Pr\left(\frac{0.08446093 - 0.05}{0.8} < \theta < \frac{0.09577916 - 0.05}{0.8}\right) &= 0.95, \\ \Pr(0.043 < \theta < 0.057) &= 0.95,\end{aligned}$$

as in (4e). With $n = 10,000$ all the approaches produce similar results.

References

Clopper, C. J. and E. S. Pearson (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika 26*, 404–413.

Diggle, P. J. (2011). Estimating prevalence using an imperfect test. *Epidemiology Research International 2011.* doi:10.1155/2011/608719.

Johnson, N. L., A. W. Kemp, and S. Kotz (2005). *Univariate Discrete Distributions*. Wiley.