Inference for Binomial Parameters

Dipankar Bandyopadhyay, Ph.D.

Department of Biostatistics,
Virginia Commonwealth University

Inference for a probability

- Phase II cancer clinical trials are usually designed to see if a new, single treatment produces favorable results (proportion of success), when compared to a known, "industry standard").
- If the new treatment produces good results, then further testing will be done in a Phase III study, in which patients will be randomized to the new treatment or the "industry standard".
- In particular, n independent patients on the study are given just one treatment, and the outcome for each patient is usually

$$Y_i = \begin{cases} 1 \text{ if new treatment shrinks tumor (success)} \\ 0 \text{ if new treatment does not shrinks tumor (failure)} \end{cases},$$

  $i = 1, \ldots, n$
- For example, suppose $n = 30$ subjects are given Polen Springs water, and the tumor shrinks in 5 subjects.
- The goal of the study is to estimate the probability of success, get a confidence interval for it, or perform a test about it.

- Suppose we are interested in testing

$$H_0 : p = .5$$

where .5 is the probability of success on the "industry standard"

As discussed in the previous lecture, there are three ML approaches we can consider.

- Wald Test (non-null standard error)
- Score Test (null standard error)
- Likelihood Ratio test

Wald Test

For the hypotheses

$$H_0 : p = p_0$$

$$H_A : p \neq p_0$$

The Wald statistic can be written as

$$\begin{aligned} z_W &= \frac{\widehat{p} - p_0}{SE} \\ &= \frac{\widehat{p} - p_0}{\sqrt{\widehat{p}(1 - \widehat{p})/n}} \end{aligned}$$

Score Test

Agresti equations 1.8 and 1.9 yield

$$u(p_0) = \frac{y}{p_0} - \frac{n-y}{1-p_0}$$

$$\iota(p_0) = \frac{n}{p_0(1-p_0)}$$

$$\begin{aligned} z_S &= \frac{u(p_0)}{[\iota(p_0)]^{1/2}} \\ &= \text{(some algebra)} \\ &= \frac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \end{aligned}$$

Application of Wald and Score Tests

- Suppose we are interested in testing

$$H_0 : p = .5,$$

- Suppose $Y = 2$ and $n = 10$ so $\widehat{p} = .2$
- Then,

$$Z_W = \frac{(.2 - .5)}{\sqrt{.2(1 - .8)/10}} = -2.37171$$

and

$$Z_S = \frac{(.2 - .5)}{\sqrt{.5(1 - .5)/10}} = -1.89737$$

- Here, $Z_W > Z_S$ and at the $\alpha = 0.05$ level, the statistical conclusion would differ.

Notes about Z_W and Z_S

- Under the null, Z_W and Z_S are both approximately $N(0,1)$. However, Z_S's sampling distribution is closer to the standard normal than Z_W so it is generally preferred.

- When testing $\text{H}_0 : p = .5$,
$$|Z_W| \geq |Z_S|$$
i.e.,
$$\left| \frac{(\widehat{p} - .5)}{\sqrt{\widehat{p}(1-\widehat{p})/n}} \right| \geq \left| \frac{(\widehat{p} - .5)}{\sqrt{.5(1-.5)/n}} \right|$$

- Why ? Note that $\widehat{p}(1-\widehat{p}) \leq .5(1-.5)$, i.e., $p(1-p)$ takes on its maximum value at $p = .5$:

p	.10	.20	.30	.40	.50	.60	.70	.80	.90
p(1-p)	.09	.16	.21	.24	.25	.24	.21	.16	.09

- Since the denominator of Z_W is always less than the denominator of $Z_S, |Z_W| \geq |Z_S|$

- Under the null, $p = .5$,

$$\widehat{p}(1-\widehat{p}) \approx .5(1-.5),$$

so

$$|Z_S| \approx |Z_W|$$

- However, under the alternative,

$$\mathrm{H}_A : p \neq .5,$$

Z_S and Z_W could be very different, and, since

$$|Z_W| \geq |Z_S|,$$

the test based on Z_W is more powerful (when testing against a null of 0.5).

- For the general test

$$H_0 : p = p_o,$$

for a specified value p_o, the two test statistics are

$$Z_S = \frac{(\widehat{p} - p_o)}{\sqrt{p_o(1 - p_o)/n}}$$

and

$$Z_W = \frac{(\widehat{p} - p_o)}{\sqrt{\widehat{p}(1 - \widehat{p})/n}}$$

- For this general test, there is no strict rule that

$$|Z_W| \geq |Z_S|$$

Likelihood-Ratio Test

- It can be shown that

$$2\log\left\{\frac{L(\widehat{p}|\mathrm{H}_A)}{L(p_o|\mathrm{H}_0)}\right\} = 2[\log L(\widehat{p}|\mathrm{H}_A) - \log L(p_o|\mathrm{H}_0)] \sim \chi^2_1$$

  where
  $L(\widehat{p}|\mathrm{H}_A)$ is the likelihood after replacing p by its estimate, $\widehat{p}$, under the alternative (H_A), and

$$L(p_o|\mathrm{H}_0)$$

  is the likelihood after replacing p by its specified value, p_o, under the null (H_0).

Likelihood Ratio for Binomial Data

- For the binomial, recall that the log-likelihood equals

$$\log L(p) = \log \left(\begin{array}{c} n \\ y \end{array} \right) + y \log p + (n-y) \log(1-p),$$

- Suppose we are interested in testing

$$\mathrm{H}_0 : p = .5 \qquad \text{versus} \qquad \mathrm{H}_0 : p \neq .5$$

- The likelihood ratio statistic generally only is for a two-sided alternative (recall it is χ^2 based)
- Under the alternative,

$$\log L(\widehat{p}|\mathrm{H}_A) = \log \left(\begin{array}{c} n \\ y \end{array} \right) + y \log \widehat{p} + (n-y) \log(1-\widehat{p}),$$

- Under the null,

$$\log L(.5|\mathrm{H}_0) = \log \left(\begin{array}{c} n \\ y \end{array} \right) + y \log .5 + (n-y) \log(1-.5),$$

Then, the likelihood ratio statistic is

$$
\begin{aligned}
2[\log L(\widehat{p}|\mathrm{H}_A) - \log L(p_o|\mathrm{H}_0)] &= 2\left[\log\begin{pmatrix} n \\ y \end{pmatrix} + y\log\widehat{p} + (n-y)\log(1-\widehat{p})\right] \\
&- 2\left[\log\begin{pmatrix} n \\ y \end{pmatrix} + y\log .5 + (n-y)\log(1-.5)\right] \\
&= 2\left[y\log\left(\tfrac{\widehat{p}}{.5}\right) + (n-y)\log\left(\tfrac{1-\widehat{p}}{1-.5}\right)\right] \\
&= 2\left[y\log\left(\tfrac{y}{.5n}\right) + (n-y)\log\left(\tfrac{n-y}{(1-.5)n}\right)\right],
\end{aligned}
$$

which is approximately χ_1^2

Example

- Recall from previous example, $Y = 2$ and $n = 10$ so $\widehat{p} = .2$
- Then, the Likelihood Ratio Statistic is

$$2\left[2\log\left(\frac{.2}{.5}\right) + (8)\log\left(\frac{.8}{.5}\right)\right] = 3.85490(p = 0.049601)$$

- Recall, both Z_W and Z_S are N(0,1), and $N(0,1)^2$ is χ_1^2
- Then, the Likelihood ratio statistic is on the same scale as Z_W^2 and Z_S^2, since both Z_W^2 and Z_S^2 are chi-square 1 df
- For this example

$$Z_S^2 = \left[\frac{(.2-.5)}{\sqrt{.5(1-.5)/10}}\right]^2 = 3.6, \text{and}$$

$$Z_W^2 = \left[\frac{(.2-.5)}{\sqrt{.2(1-.8)/10}}\right]^2 = 5.625$$

- The Likelihood Ratio Statistic is between Z_S^2 and Z_W^2.

Likelihood Ratio Statistic

For the general test

$$\text{H}_0 : p = p_o,$$

the Likelihood Ratio Statistic is

$$2\left[y\log\left(\frac{\widehat{p}}{p_o}\right) + (n-y)\log\left(\frac{1-\widehat{p}}{1-p_o}\right)\right] \sim \chi_1^2$$

asymptotically under the Null.

Large Sample Confidence Intervals

- In large samples, since

$$\widehat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) ,$$

  we can obtain a 95% confidence interval for p with

$$\widehat{p} \pm 1.96\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$$

- However, since $0 \leq p \leq 1$, we would want the endpoints of the confidence interval to be in $[0, 1]$, but the endpoints of this confidence interval are not restricted to be in $[0, 1]$.
- When p is close to 0 or 1 (so that $\widehat{p}$ will usually be close to 0 or 1), and/or in small samples, we could get endpoints outside of [0,1]. The solution would be the truncate the interval endpoint at 0 or 1.

Example

- Suppose $n = 10$, and $Y = 1$, then

$$\widehat{p} = \frac{1}{10} = .1$$

and the 95% confidence interval is

$$\widehat{p} \pm 1.96\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},$$

$$.1 \pm 1.96\sqrt{\frac{.1(1-.1)}{10}},$$

$$[-.086, .2867]$$

- After truncating, you get,

$$[0, .2867]$$

Exact Test Statistics and Confidence Intervals

Unfortunately, many of the phase II trials have small samples, and the above asymptotic test statistics and confidence intervals have very poor properties in small samples. (A 95% confidence interval may only have 80% coverage). In this situation, "Exact test statistics and Confidence Intervals" can be obtained.

One-sided Exact Test Statistic

- The historical norm for the clinical trial you are doing is 50%, so you want to test if the response rate of the new treatment is greater then 50%.
- In general, you want to test

$$\mathrm{H}_0{:}p = p_o = 0.5$$

versus

$$\mathrm{H}_A{:}p > p_o = 0.5$$

- The test statistic

$$Y = \text{the number of successes out of } n \text{ trials}$$

Suppose you observe y_{obs} successes ;

Under the null hypothesis,

$$n\widehat{p} = Y \sim Bin(n, p_o),$$

i.e.,

$$P(Y = y|\mathrm{H}_0{:}p = p_o) = \left(\begin{array}{c} n \\ y \end{array} \right) p_o^y (1 - p_o)^{n-y}$$

- When would you tend to reject $\mathrm{H}_0{:}p = p_o$ in favor of $\mathrm{H}_A{:}p > p_o$

Answer

Under H_0:$p = p_o$, you would expect $\widehat{p} \approx p_o$
$(Y \approx np_o)$
Under H_A:$p > p_o$, you would expect $\widehat{p} > p_o$
$(Y > np_o)$
i.e., you would expect Y to be 'large' under the alternative.

Exact one-sided p-value

- If you observe y_{obs} successes, the exact one-sided p-value is the probability of getting the observed y_{obs} plus any larger (more extreme) Y

$$\begin{aligned} p-\text{value} &= \mathrm{pr}(Y \geq y_{obs} | \mathrm{H}_0 : p = p_o) \\ &= \sum_{j=y_{obs}}^{n} \binom{n}{j} p_o^j (1-p_o)^{n-j} \end{aligned}$$

Other one-sided exact p-value

- You want to test

$$H_0{:}p = p_o$$

versus

$$H_A{:}p < p_o$$

- The exact p-value is the probability of getting the observed y_{obs} plus any smaller (more extreme) y

$$\begin{aligned} p - \text{value} &= \mathrm{pr}(Y \leq y_{obs} | \mathrm{H}_0{:}p = p_o) \\ &= \sum_{j=0}^{y_{obs}} \binom{n}{j} p_o^j (1 - p_o)^{n-j} \end{aligned}$$

Two-sided exact p-value

- The general definition of a 2-sided exact p-value is

$$P\left[\begin{array}{l} \text{seeing a result as likely or} \\ \text{less likely than the observed result} \end{array} \middle| \mathrm{H}_0 \right].$$

It is easy to calculate a 2-sided $p-$value for a symmetric distribution, such as $Z \sim N(0,1)$. Suppose you observe $z > 0$,

```
fontsize=
fontsize=
fontsize= 0.4 +                         ...     Standard Normal Density
fontsize=     |                       ... ...
fontsize=     |                       .     .
fontsize=     |                      ..     ..
fontsize=              ******** Graph Modified to Fit on Slide *********
fontsize=     |                 ..                    ..
fontsize=     |                 .                     .
fontsize=     |                ..                     ..
fontsize= 0.1 +      less likely  .|                      |.   less likely
fontsize=     |            <==== ..|                      |..  ====>
fontsize=     |                 .. |                      | ..
fontsize=     |                ..  |                      |  ..
fontsize=     |              ...   |                      |   ...
fontsize=     |           ....     |                      |     ....
fontsize= 0.0 +  ..........        |                      |        ..........
fontsize=     |                    |                      |
fontsize=     ---+-------------+--|-----------+-----------|--+-------------+--
fontsize=       -4            -2  |           0           |  2             4
fontsize=                         |                       |
fontsize=                       -1.96                    1.96
fontsize=                         -z                      z
fontsize=
```

Symmetric distributions

- If the distribution is symmetric with mean 0, e.g., normal, then the exact 2-sided $p-$value is

$$p - \text{value} = 2 \cdot P(Z \geq |z|)$$

  when z is positive or negative.

- In general, if the distribution is symmetric, but not necessarily centered at 0, then the exact 2-sided $p-$value is

$$p - \text{value} = 2 \cdot \ \min\{P(Y \geq y_{obs}), P(Y \leq y_{obs})\}$$

- Now, consider a symmetric binomial. For example, suppose $n = 4$ and $p_o = .5$, then,

```
Binomial PDF for N=4 and P=0.5

Number of
 Successes      P(Y=y)     P(Y<=y)     P(Y>=y)

     0          0.0625      0.0625      1.0000
     1          0.2500      0.3125      0.9375
     2          0.3750      0.6875      0.6875
     3          0.2500      0.9375      0.3125
     4          0.0625      1.0000      0.0625
```

Suppose you observed $y_{obs} = 4$, then the exact two-sided p-value would be

$$
\begin{aligned}
p - \text{value} &= 2 \cdot \min\{\mathrm{pr}(Y \geq y_{obs}), \mathrm{pr}(Y \leq y_{obs})\} \\
&= 2 \cdot \min\{\mathrm{pr}(Y \geq 4), \mathrm{pr}(Y \leq 4)\} \\
&= 2 \cdot \min\{.0625, 1\} \\
&= 2(.0625) \\
&= .125
\end{aligned}
$$

- The two-sided exact p-value is trickier when the binomial distribution is not symmetric
- For the binomial data, the exact 2-sided p-value is

$$P\left[\begin{array}{l}\text{seeing a result as likely or}\\ \text{less likely than the observed}\\ \text{result in either direction}\end{array}\,\middle|\, \mathrm{H}_0 : p = p_o\right].$$

- Essentially the sum of all probabilities such that $P(Y = y|P_0) \leq P(y_{obs}|P_0)$

In general, to calculate the 2-sided $p-$value

1. Calculate the probability of the observed result under the null

$$\pi = P(Y = y_{obs} | p = p_o) = \begin{pmatrix} n \\ y_{obs} \end{pmatrix} p_o^{y_{obs}} (1 - p_o)^{n - y_{obs}}$$

2. Calculate the probabilities of all $n + 1$ values that Y can take on:

$$\pi_j = P(Y = j | p = p_o) = \begin{pmatrix} n \\ j \end{pmatrix} p_o^j (1 - p_o)^{n-j},$$

   $j = 0, \ldots, n.$

3. Sum the probabilities π_j in (2.) that are less than or equal to the observed probability π in (1.)

$$p - value = \sum_{j=0}^{n} \pi_j I(\pi_j \leq \pi) \text{ where}$$

$$I(\pi_j \leq \pi) = \begin{cases} 1 \text{ if } \pi_j \leq \pi \\ 0 \text{ if } \pi_j > \pi \end{cases} .$$

- Suppose $n = 5$, you hypothesize $p = .4$ and we observe $y = 3$ successes.
- Then, the PDF for this binomial is

```
Binomial PDF for N=5 and P=0.4

Number of
 Successes      P(Y=y)     P(Y<=y)     P(Y>=y)

     0         0.07776     0.07776     1.00000
     1         0.25920     0.33696     0.92224
     2         0.34560     0.68256     0.66304
     3         0.23040     0.91296     0.31744   <----Y obs
     4         0.07680     0.98976     0.08704
     5         0.01024     1.00000     0.01024
```

Exact p-value by hand

- Step 1: Determine $P(Y = 3|n = 5, P_0 = .4)$. In this case $P(Y = 3) = .2304$.
- Step 2: Calculate Table (see previous slide)
- Step 3: Sum probabilities less than or equal to the one observed in step 1. When $Y \in \{0, 3, 4, 5\}$, $P(Y) \leq 0.2304$.

ALTERNATIVE	EXACT	PROBS
H_A: $p > .4$	.317	$P[Y \geq 3]$
H_A: $p < .4$	.913	$P[Y \leq 3]$
H_A: $p \neq .4$	.395	$P[Y \geq 3] + P[Y = 0]$

Comparison to Large Sample Inference

Note that the exact and asymptotic do not agree very well:

ALTERNATIVE	EXACT	LARGE SAMPLE
H_A: $p > .4$	.317	.181
H_A: $p < .4$	.913	.819
H_A: $p \neq .4$	.395	.361

We will look at calculations by

1. `STATA` (best)
2. `R` (good)
3. `SAS` (surprisingly poor)

The following STATA code will calculate the exact p-value for you

```
From within STATA at the dot, type

bitesti 5 3 .4


----------Output-------------------------------------------
        N   Observed k   Expected k   Assumed p   Observed p
------------------------------------------------------------
        5            3            2     0.40000      0.60000

  Pr(k >= 3)            = 0.317440  (one-sided test)
  Pr(k <= 3)            = 0.912960  (one-sided test)
  Pr(k <= 0 or k >= 3) = 0.395200  (two-sided test)
```

To perform an exact binomial test in R
Use the `binom.test` function available in R package `stats`

```
> binom.test(3, 5, p = 0.4, alternative = "two.sided")
        Exact binomial test
data:  3 and 5
number of successes = 3, number of trials = 5, p-value = 0.3952
alternative hypothesis: true probability of success is not equal to
95 percent confidence interval:
 0.1466328 0.9472550
sample estimates: probability of success
                   0.6
```

This gets a score of good since the output is not as descriptive as the STATA output.

Interestingly, `SAS Proc Freq` gives the wrong 2-sided $p-$value

```
data one;
 input outcome $ count;
cards;
1succ 3
2fail 2
;
proc freq data=one;
 tables outcome / binomial(p=.4);
 weight count;
 exact binomial;
run;
```

```
---------Output-----------

        Binomial Proportion
        for outcome = 1succ
-----------------------------------
   Test of H0: Proportion = 0.4

ASE under H0                 0.2191
Z                            0.9129
One-sided Pr >  Z            0.1807
Two-sided Pr > |Z|           0.3613

Exact Test
One-sided Pr >=  P           0.3174
Two-sided = 2 * One-sided    0.6349

Sample Size = 5
```

Better Approximation using the normal distribution

- Because Y is discrete, a 'continuity-correction' is often applied to the normal approximation to more closely approximate the exact $p-$value.
- To make a discrete distribution look more approximately continuous, the probability function is drawn such that $\mathrm{pr}(Y = y)$ is a rectangle centered at y with width 1, and height $\mathrm{pr}(Y = y)$, i.e.,
- The area under the curve between $y - 0.5$ and $y + 0.5$ equals

$$[(y + 0.5) - (y - 0.5)] \cdot P(Y = y) = 1 \cdot P(Y = y)$$

For example, suppose as before, we have $n = 5$ and $p_o = .4,.$

Then on the probability curve,

$$\mathrm{pr}(Y \geq y) \approx \mathrm{pr}(Y \geq y - .5)$$

which, using the continuity corrected normal approximation is

$$\mathrm{pr}\left(Z \geq \frac{(y - .5) - np_o}{\sqrt{np_o(1 - p_o)}} \middle| \mathrm{H}_0{:}p = p_o; Z \sim N(0,1) \right)$$

and

$$\mathrm{pr}(Y \leq y) \approx \mathrm{pr}(Y \leq y + .5)$$

which, using the continuity corrected normal approximation

$$\mathrm{pr}\left(Z \leq \frac{(y + .5) - np_o}{\sqrt{np_o(1 - p_o)}} \middle| \mathrm{H}_0{:}p = p_o; Z \sim N(0,1) \right)$$

With the continuity correction, the above $p-$values becomes

ALTERNATIVE	EXACT	LARGE SAMPLE	Continuity Corrected LARGE SAMPLE
H_A: $p > .4$	.317	.181	.324
H_A: $p < .4$	.913	.819	.915
H_A: $p \neq .4$	.395	.361	.409

Then, even with the small sample size of $n = 5$, the continuity correction does a good job of approximating the exact $p-$value.

Also, as $n \to \infty$, the exact and asymptotic are equivalent under the null; so for large n, you might as well use the asymptotic.

However, given the computational power available, you can easily calculate the exact p-value.

Exact Confidence Interval

A $(1-\alpha)$ confidence interval for p is of the form

$$[p_L, p_U],$$

where p_L and p_U are random variables such that

$$\mathrm{pr}[p_L \leq p \leq p_U] = 1-\alpha$$

For example, for a large sample 95% confidence interval,

$$p_L = \widehat{p} - 1.96\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},$$

and

$$p_U = \widehat{p} + 1.96\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},$$

.

It can be shown that, to obtain a 95% exact confidence interval $[p_L, p_U]$, the endpoints p_L and p_U satisfy

$$\begin{aligned}
\alpha/2 = .025 &= \mathrm{pr}(Y \geq y_{obs} | p = p_L) \\
&= \sum_{j=y_{obs}}^{n} \binom{n}{j} p_L^j (1-p_L)^{n-j},
\end{aligned}$$

and

$$\begin{aligned}
\alpha/2 = .025 &= \mathrm{pr}(Y \leq y_{obs} | p = p_U) \\
&= \sum_{j=0}^{y_{obs}} \binom{n}{j} p_U^j (1-p_U))^{n-j}
\end{aligned}$$

- In these formulas, we know $\alpha/2 = .025$ and we know y_{obs} and n. Then, we solve for the unknowns p_L and p_U.
- Can figure out p_L and p_U by plugging different values for p_L and p_U until we find the values that make $\alpha/2 = .025$
- Luckily, this is implemented on the computer, so we don't have to do it by hand.
- Because of relationship between hypothesis testing and confidence intervals, to calculate the exact confidence interval, we are actually setting the exact one-sided $p-$values to $\alpha/2$ for testing $\mathrm{H}_o : p = p_o$ and solving for p_L and p_U.
- In particular, we find p_L and p_U to make these $p-$values equal to $\alpha/2$.

Example

- Suppose $n = 5$ and $y_{obs} = 4$, and we want a 95% confidence interval. ($\alpha = .05$, $\alpha/2 = .025$).
- Then, the lower point, p_L of the exact confidence interval $[p_L, p_U]$ is the value p_L such that

$$\alpha/2 = .025 = \mathrm{pr}[Y \geq 4|p = p_L] = \sum_{j=4}^{5} \left(\begin{array}{c} 5 \\ j \end{array} \right) p_L^j (1 - p_L)^{n-j},$$

- If you don't have a computer program to do this, you can try 'trial' and error for p_L

p_L	$\mathrm{pr}(Y \geq 4 \mid p = p_L)$
.240	0.013404
.275	0.022305
.2836	$.025006^* \approx .025$

- Then, $p_L \approx .2836$.

- Similarly, the upper point, p_U of the exact confidence interval $[p_L, p_U]$ is the value p_U such that

$$\alpha/2 = .025 = \mathrm{pr}[Y \le 4 | p = p_U] = \sum_{j=0}^{4} \left(\begin{array}{c} 5 \\ j \end{array} \right) p_U^j (1 - p_U)^{n-j},$$

- Similarly, you can try "trial" and error for the p_U

p_U	$\mathrm{pr}(Y \le 4 \vert p = p_U)$
.95	0.22622
.99	0.049010
.994944	$0.025026^* \approx .025$

STATA ? The following STATA code will calculate the exact binomial confidence interval for you

```
. cii 5 4
----------- Output -----------------
                                           -- Binomial Exact --
Variable |      Obs      Mean    Std. Err.    [95% Conf. Interval]
---------+-----------------------------------------------------------------
         |        5        .8    .1788854      .2835937    .9949219
```

How about SAS?

```
data one;
 input outcome $ count;
cards;
1succ 4
2fail 1
;
proc freq data=one;
 tables outcome / binomial;
 weight count;
run;
```

```
      Binomial Proportion
--------------------------------
Proportion                0.8000
ASE                       0.1789
95% Lower Conf Limit      0.4494
95% Upper Conf Limit      1.0000

Exact Conf Limits
95% Lower Conf Limit      0.2836
95% Upper Conf Limit      0.9949

  Test of H0: Proportion = 0.5

ASE under H0              0.2236
Z                         1.3416
One-sided Pr >  Z         0.0899
Two-sided Pr > |Z|        0.1797
```

Comparing the exact and large sample

- Then, the two sided confidence intervals are

EXACT	LARGE SAMPLE (NORMAL) $\widehat{p}$
[.2836,.9949]	[.449,1]

- We had to truncate the upper limit based on using $\widehat{p}$ at 1.
- The exact CI is not symmetric about $\widehat{p} = \frac{4}{5} = .8$, whereas the the confidence interval based on $\widehat{p}$ would be if not truncated.
- Suggestion; if $Y < 5$, and/or $n < 30$, use exact; for large Y and n, any of the three would be almost identical.

Exact limits based on F Distribution

- While software would be the tool of choice (I doubt anyone still calculates exact binomial confidence limits by hand), there is a distributional relationship among the Binomial and F distributions.
- In particular P_L and P_U can be found using the following formulae

$$P_L = \frac{y_{obs}}{y_{obs} + (n - y_{obs} + 1)F_{2(n-y_{obs}+1), 2\cdot y_{obs}, 1-\alpha/2}}$$

- and

$$P_U = \frac{(y_{obs} + 1) \cdot F_{2\cdot(y_{obs}+1), 2\cdot(n-y_{obs}), 1-\alpha/2}}{(n - y_{obs}) + (y_{obs} + 1) \cdot F_{2\cdot(y_{obs}+1), 2\cdot(n-y_{obs}), 1-\alpha/2}}$$

Example using F-dist

- Thus, using our example of $n = 5$ and $y_{obs} = 4$

$$\begin{aligned} P_L &= \frac{y_{obs}}{y_{obs} + (n - y_{obs} + 1) F_{2(n - y_{obs} + 1), 2 \cdot y_{obs}, 1 - \alpha/2}} \\ &= \frac{4}{4 + 2F_{4,8,0.975}} \\ &= \frac{4}{4 + 2 \cdot 5.0526} \\ &= 0.2836 \end{aligned}$$

- and

$$\begin{aligned} P_U &= \frac{(y_{obs} + 1) \cdot F_{2 \cdot (y_{obs} + 1), 2 \cdot (n - y_{obs}), 1 - \alpha/2}}{(n - y_{obs}) + (y_{obs} + 1) \cdot F_{2 \cdot (y_{obs} + 1), 2 \cdot (n - y_{obs}), 1 - \alpha/2}} \\ &= \frac{5 \cdot F_{10,2,0.975}}{1 + 5 \cdot F_{10,2,0.975}} \\ &= \frac{5 \cdot 39.39797}{1 + 5 \cdot 39.39797} \\ &= 0.9949 \end{aligned}$$

- Therefore, our 95% exact confidence interval for p is $[0.2836, 0.9949]$ as was observed previously

```
%macro mybinomialpdf(p,n);
dm "output" clear; dm "log" clear;
options nodate nocenter nonumber;
data myexample;
  do i = 0 to &n;
    prob = PDF('BINOMIAL',i,&p,&n) ;
    cdf = CDF('BINOMIAL',i,&p,&n)  ;
    m1cdfprob = 1-cdf+prob;
    output;
  end;
  label i = "Number of *Successes";
  label prob =    "P(Y=y)  ";
  label cdf =     "P(Y<=y)";
  label m1cdfprob="P(Y>=y)";
run;
title "Binomial PDF for N=&n and P=&p";
proc print noobs label split="*";
run;
%mend mybinomialpdf;
%mybinomialpdf(0.4,5);
```

1.4.3 where for art thou, vegetarians?

Out of $n = 25$ students, $y = 0$ were vegetarians. Assuming binomial data, the 95% CIs found by inverting the Wald, score, and LRT tests are

Wald	$(0, 0)$
score	$(0, 0.133)$
LRT	$(0, 0.074)$

The Wald interval is particularly troublesome. Why the difference? for small or large (true, unknown) π the normal approximation for the distribution of $\hat{\pi}$ is pretty bad in small samples.

A solution is to consider the *exact* sampling distribution of $\hat{\pi}$ rather than a normal approximation.

1.4.4 Exact inference

An exact test proceeds as follows.

Under $H_0 : \pi = \pi_0$ we know $Y \sim \text{bin}(n, \pi_0)$. Values of $\hat{\pi}$ far away from π_0, or equivalently, values of Y far away from $n\pi_0$, indicate that $H_0 : \pi = \pi_0$ is unlikely.

Say we reject H_0 if $Y < a$ or $Y > b$ where $0 \leq a < b \leq n$. Then we set the type I error at α by requiring $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha$. That is,

$$P(Y < a | \pi = \pi_0) = \frac{\alpha}{2} \text{ and } P(Y > b | \pi = \pi_0) = \frac{\alpha}{2}.$$

Bounding Type I error

However, since Y is discrete, the best we can do is *bounding* the type I error by choosing a as large as possible such that

$$P(Y < a|\pi = \pi_0) = \sum_{i=0}^{a-1} \binom{n}{i} \pi_0^i (1-\pi_0)^{n-i} < \frac{\alpha}{2},$$

and b as small as possible such that

$$P(Y > b|\pi = \pi_0) = \sum_{i=b+1}^{n} \binom{n}{i} \pi_0^i (1-\pi_0)^{n-i} < \frac{\alpha}{2}.$$

Exact test, cont.

For example, when $n = 20$, $H_0 : \pi = 0.25$, and $\alpha = 0.05$ we have

$$P(Y < 2|\pi = 0.25) = 0.024 \text{ and } P(Y < 3|\pi = 0.25) = 0.091,$$

so $a = 2$. Also,

$$P(Y > 9|\pi = 0.25) = 0.014 \text{ and } P(Y > 8|\pi = 0.25) = 0.041,$$

so $b = 9$. We reject $H_0 : \pi = 0.25$ when $Y < 2$ or $Y > 9$. The type I error is bounded: $\alpha = P(\text{reject } H_0|H_0 \text{ is true}) \leq 0.05$, but in fact this is conservative, $P(\text{reject } H_0|H_0 \text{ is true}) = 0.024 + 0.014 = 0.038$.

Nonetheless, this type of exact test can be inverted to obtain exact confidence intervals for π. However, the actual coverage probability is *at least* as large as $1 - \alpha$, but typically more. So the procedure errs on the side of being conservative (CI's are bigger than they need to be). Section 16.6.1 has more details.

Tests in R

To obtain the 95% CI from inverting the score test, and from inverting the exact (Clopper-Pearson) test:

```
> out1=prop.test(x=0,n=25,conf.level=0.95,correct=F)
> out1$conf.int
[1] 0.0000000 0.1331923
attr(,"conf.level") [1] 0.95
> out2=binom.test(x=0,n=25,conf.level=0.95)
> out2$conf.int
[1] 0.0000000 0.1371852
attr(,"conf.level") [1] 0.95
```

SAS code

```
data table;
 input vegetarian$ count @@;
 datalines;
 yes 0 no 25
;
* let pi be proportion of vegetarians in population;
* let's test H0: pi=0.032 (U.S. proportion) and obtain exact 95% CI for pi;
* SAS also provides a test of H0: pi=0.5,
* other options given by binomial(ac wilson exact jeffreys)
* even though you didn't ask for it! (not shown on next slide);
proc freq data=table order=data; weight count / zeros;
 tables vegetarian / binomial testp=(0.032,0.968);
 exact binomial chisq;
run;
data veg;
input response $ count;
datalines;
no 25
yes 0
;
proc freq data=veg; weight count;
tables response / binomial(ac wilson exact jeffreys) alpha
=.05; run;
```

SAS output

```
                    The FREQ Procedure

                                         Test     Cumulative    Cumulative
vegetarian    Frequency     Percent    Percent     Frequency      Percent
---------------------------------------------------------------------------
yes                   0        0.00       3.20             0         0.00
no                   25      100.00      96.80            25       100.00

                    Chi-Square Test
               for Specified Proportions
          ---------------------------------------
          Chi-Square                      0.8264
          DF                                   1
          Asymptotic Pr >  ChiSq          0.3633
          Exact      Pr >= ChiSq          0.6335

  WARNING: 50% of the cells have expected counts less than 5.
           (Asymptotic) Chi-Square may not be a valid test.

                   Binomial Proportion
                 for vegetarian = yes
           -------------------------------------
           Proportion (P)                0.0000
           ASE                           0.0000
           95% Lower Conf Limit          0.0000
           95% Upper Conf Limit          0.0000
           Exact Conf Limits
           95% Lower Conf Limit          0.0000
           95% Upper Conf Limit          0.1372
                      Sample Size = 25
```