3. Analyses of Binary Outcome Measures

3.1 Treatment Effect for Binary Outcome Measures

Suppose the outcome measure Y_i is binary, examples of which might include death, survival, recurrence or remission from disease, sometimes referred to by the neutral term "event". One summary of outcome is the proportion of patients that had the event in each treatment group, which estimates the probability of events in each treatment, say π_T and π_C, or population rates. An alternative summary statistics is the odds of the event, which is the probability of the event divided by the probability of the complimentary event.

Table 3.1 Notation for a Trial with a Binary Outcome

Frequency Dist.	Treatment	Control
Yes	r_T	r_C
No	$n_T - r_T$	$n_C - r_C$
Total	n_T	n_C
Probability of Event (Population proportion)	π_T	π_C
Sample proportion	$p_T = \frac{r_T}{n_T}$	$p_C = \frac{r_C}{n_C}$
Odds of Event Population Odds	$\frac{\pi_T}{(1-\pi_T)}$	$\frac{\pi_C}{(1-\pi_C)}$
Sample Odds	$q_T = \frac{r_T}{n_T - r_T}$	$q_C = \frac{r_C}{n_C - r_C}$

The effect of treatment can be measured in three ways

Rate or Risk Difference, $RD = \pi_T - \pi_C$,

Rate Ratio, $RR = \dfrac{\pi_T}{\pi_C}$

Odd Ratio, $OR = \dfrac{\dfrac{\pi_T}{(1-\pi_T)}}{\dfrac{\pi_C}{(1-\pi_C)}} = \dfrac{\pi_T(1-\pi_C)}{(1-\pi_T)\pi_C}$

3.2 Inference for the Rate Difference

The rate difference (*RD*) is estimated by $\hat{RD} = p_T - p_C$ where

$p_T = \dfrac{r_T}{n_T}$ and $p_C = \dfrac{r_C}{n_C}$. The numbers of successes r_T and r_C have

distributions $Bin[n_T, \pi_T]$ and $Bin[n_C, \pi_C]$. From properties of the

binomial distribution the variance of r equals $n.\pi(1-\pi)$. Hence the

proportion $p = \dfrac{r}{n}$ is given by $Var\left[\dfrac{r}{n}\right] = \dfrac{Var[r]}{n^2} = \dfrac{\pi(1-\pi)}{n}$.

Since treatment groups are independent, it follows that

$$Var\left[\hat{RD}\right] = Var[p_T - p_C] = Var[p_T] + Var[p_C] = \frac{\pi_T(1-\pi_T)}{n_T} + \frac{\pi_C(1-\pi_C)}{n_C}$$

.

This can be estimated by substituting p_T and p_C for π_T and π_C.

Hence,

$$\hat{SE}\left[\hat{RD}\right] = \hat{SE}[p_T - p_C] = \sqrt{\frac{p_T(1-p_T)}{n_T} + \frac{p_C(1-p_C)}{n_C}},$$

which is used for confidence interval construction.

Under the null hypothesis $H_0 : RD = 0$, $\pi_T = \pi_C = \pi$ say. The pooled

proportion π can be estimated by $p = \dfrac{r_T + r_C}{n_T + n_C}$. Hence, the null

standard error can be defined as

$$SE_{null}\left[\widehat{RD}\right] = SE[p - p] = \sqrt{\left(\frac{p(1-p)}{n_T} + \frac{p(1-p)}{n_C}\right)} = \sqrt{p(1-p)\left(\frac{1}{n_T} + \frac{1}{n_C}\right)}$$

which is used for statistical inference on $H_0 : RD = 0$.

Two-Sample z-test of Proportions

A test of $H_0: RD = 0$ vs $H_1: RD \neq 0$ can be constructed as

$$Z_{RD} = \frac{\hat{RD}}{\hat{SE}_{null}\left[\hat{RD}\right]} = \frac{p_T - p_C}{\hat{SE}_{null}\left[p_T - p_C\right]}.$$

Under assumptions given below Z_{RD} is approximates a standardised normal distribution, $N[0,1]$. This gives the two-sample z-test for proportions corresponding to the two-sample t-test of means.

Two-Sample z-test of Proportions

$$Z_{RD} = \frac{p_T - p_C}{\hat{SE}_{null}\left[p_T - p_C\right]} \text{ and } \hat{SE}_{null}\left[p_T - p_C\right] = \sqrt{p(1-p)\left(\frac{1}{n_T} + \frac{1}{n_C}\right)}$$

where $p = \frac{r_T + r_C}{n_T + n_C}$.

For an α-size two-sided test of $H_0: RD = 0$ vs $H_1: RD \neq 0$ compare Z_{RD} against critical values defined by $\pm z_{\alpha/2}$. Alternatively, the p-values for the two-sided test is given by $2\left(1 - \Phi\left(|Z_{RD}|\right)\right)$ where Φ is the cumulative density of the standardized normal distribution.

Confidence Interval

A (1-α) confidence interval for $RD = \pi_T - \pi_C$ is given by

$$p_T - p_C \pm z_{\alpha/2}\hat{SE}\left[p_T - p_C\right]$$

where $\hat{SE}\left[p_T - p_C\right] = \sqrt{\frac{p_T(1-p_T)}{n_T} + \frac{p_C(1-p_C)}{n_C}}$

Assumptions

(i) subjects are independent and

(ii) $n_T\,p$, $n_C\,p$, $n_T\,(1\text{-}p)$, $n_C\,(1\text{-}p)$ are all greater than 5.

There are improved formulae for the z-test and confidence interval that include a *continuity correction* to improve the normal approximation of the binomial distribution, but these methods are not considered in the module.

Example 3.1. The Propranolol Trial

91 patients admitted with myocardial infarction were randomly allocated to propranolol or placebo. The table below records survival status of propranolol treated patients and control patients 28 days after admission

Status 28 days after admission	*Propranolol*	*Placebo*	Total
Alive	r_T 38 (84%)	r_C 29 (63%)	67
Died	7	17	24
Total	n_T 45	n_C 46	91

Ex 3.1 *For the Propranolol Trial data calculates the point estimate of the difference in survival rate*

For the Propranolol group the rate = $\frac{38}{45} = 0.84$ (84%)

For the Placebo group the rate = $\frac{29}{46} = 0.63$ (63%)

Therefore RD= $0.84 - 0.63 = 0.21$ (21%)

Ex 3.2 *Check the assumptions of z-test of proportions*

The assumptions of the z-test of proportions are that $n_T\,p$, $n_C\,p$, $n_T\,(1\text{-}p)$, $n_C\,(1\text{-}p)$ are all greater than 5.

$$p = \frac{r_T + r_C}{n_T + n_C} = \frac{67}{91}$$

Hence $n_T\,p$, $n_C\,p$, $n_T\,(1\text{-}p)$, $n_C\,(1\text{-}p)$ are

The smallest will be for the smallest column and row total, that $n_T(1-p)$

$n_T(1-p) = 45 \times \frac{67}{91} = 11.88 > 5$!

Assumption for normal approximation satisfied.

Ex 3.3 *Compare the survival rate for the two treatments using a z-test of proportions.*

From above p=0.736. $p_T - p_C = \frac{r_T}{n_T} - \frac{r_C}{n_C} = \frac{38}{45} - \frac{29}{46} = 0.214$

$$\hat{SE}_{null}[p_T - p_C] = \sqrt{p(1-p)\left(\frac{1}{n_T} + \frac{1}{n_C}\right)} = \sqrt{0.736 \times 0.264 \left(\frac{1}{45} + \frac{1}{46}\right)} = 0.0924$$

$$Z_{RD} = \frac{p_T - p_C}{\hat{SE}_{null}[p_T - p_C]} = \frac{0.214}{0.0924} = 2.316$$

From table of the normal distribution the critical values of a two-sided 5% level test are ± 1.960

$$p - value = 2\left(1 - \Phi\left(|Z|\right)\right) = 2(1 - 0.9896) = 0.0208$$

Note that it does not matter whether the z-test is compute based on the proportion who have died or the proportion still alive.

Normal Distribution and Normal Statistical Tables

Suppose Φ is the cumulative distribution function of a standardized normal distribution $N[0,1]$. In this module the percentage point z_α of a random variable Z with distribution $N[0,1]$ is the value such that $P[Z > z_\alpha] = 1 - \Phi(z_\alpha) = \alpha$.

Tables provided by the Mathematic department define a percentage point the percentage point z_q to be the value such that

$P[Z < z_q] = \Phi(z_q) = q$.

Table 3.2 Summary of Percentage Points of the Standardized Normal Distribution

α	q	z
0.2	0.8	0.8416
0.1	0.9	1.2816
0.05	0.95	1.6449
0.025	0.975	1.9600
0.01	0.99	2.3263
0.005	0.995	2.5758

Calculation of 95%-Confidence interval for difference of proportions

Ex 3.4 *For the Propranolol Trial calculate a 95% confidence interval of the difference in survival rate for the two treatments.*

$$\hat{SE}[p_T - p_C] = \sqrt{\frac{p_T(1-p_T)}{n_T} + \frac{p_C(1-p_C)}{n_C}} = \sqrt{\frac{0.16 \times 0.84}{45} + \frac{0.37 \times 0.63}{46}} = 0.0893$$

From tables $z_{0.025} = 1.96$

(1-α) confidence interval calculated from $p_T - p_C \pm z_{\alpha/2}\hat{SE}[p_T - p_C]$

$0.214 \pm 1.96 \times 0.0893$

Hence 95% c.i. is

0.03887 to 0.3891

Ex 3.5 *Briefly comment on the effect of propranolol treatment on survival.*

" There was evidence that for patients admitted with myocardial infarction those treated with propranolol had an improved survival at 28 days post admission (84%) as compared to untreated patients (63%) with a difference of 21% (95% c.i. 3.9% to 38.9% p= 0.0208)."

Note. In the above statement p refers to the p-value. In the notes above it refers to the sample proportion under the null hypothesis.

Figure 3.1 STATA Output for z-test of proportions for the Propranolol Trial Data based on numbers of death before 28 days

```
Two-sample test of proportion                 Placebo: Number of obs =       46
                                              Propranolol: Number of obs =   45
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Placebo |   .3695652   .0711683                      .2300779    .5090526
  Propranolol |   .1555556   .0540284                      .0496619    .2614493
-------------+----------------------------------------------------------------
        diff |   .2140097   .0893532                      .0388806    .3891387
             |  under Ho:   .0923926     2.32   0.021
------------------------------------------------------------------------------

     diff = prop(Placebo) - prop(Propranolol)                      z =   2.3163
                                   Ho: diff = 0
      Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
  Pr(Z < z) = 0.9897         Pr(|Z| < |z|) = 0.0205          Pr(Z > z) = 0.0103
```

One sided lower Two sided One sided upper

Numbers Need to Treat

Numbers need to treat (NNT) is defined as the average of the number of patients that need to be treated to prevent one additional bad outcome. This measure is popular with doctors as it gives them a measure of the population level benefit of any treatment.

NNT is simply the reciprocal of the rate difference RD, that

$$NNT = \frac{1}{RD}.$$

A confidence interval of NTT can be found by taking the reciprocal of the confidence limits of *RD*. Note that the confidence interval becomes nonsensical if the confidence interval of RD includes zero.

Ex 3.6 Calculate the point estimate and 95% confidence interval of NNT for propranolol.

NNT = 1/0.214 = 4.67

95% C.I. of NNT is 1/0.03887 to 1/0.3991

NNT is 4.7 (95% c.i. 2.6 to 25.7)

Confidence interval of NNT become complicated when interval for RD includes zero

3.3 Inference Based on the Odd Ratio

The odd ratio $OR = \frac{\pi_T(1-\pi_C)}{(1-\pi_T)\pi_C}$ can be estimated by $\hat{OR} = \frac{p_T/(1-p_T)}{p_C/(1-p_C)}$.

Since $p_T = \frac{r_T}{n_T}$ and $p_C = \frac{r_C}{n_C}$, $\hat{OR} = \frac{r_T(n_C - r_C)}{(n_T - r_T)r_C}$.

Example cont. Propranolol Trial

Status 28 days after admission	***Propranolol***	***Placebo***
Alive	38	29
Died	7	17
Total	45	46
Proportion surviving	84% (38/45)	63% (29/46)

Ex 3.7 Calculate the odds ratio of survival until 28 days for propranolol treatment as compared to placebo

$$\hat{OR} = \frac{r_T(n_C - r_C)}{(n_T - r_T)r_C} = \frac{38 \times 17}{7 \times 29} = 3.1823$$

The odds ratio takes values in the range $(0,\infty)$. An odds ratio equal to 1 implies no effect. If the odds ratio is greater than 1, it implies increased odds and below 1 implies reduced odds. The odd ratio for an event (say death) is the reciprocal of the odd ratio for the complimentary event (say survival).

Confidence Intervals for Odds Ratios

The sampling distribution of odds ratio (OR) is poorly approximated by the normal distribution. Instead, confidence intervals for the $\log_e[OR]$ are calculated and then exponents (anti-logs) taken to get the confidence interval of the odds ratio. This means that the resulting confidence interval is not symmetric about the point estimate.

$$SE\left[\log_e\left[\hat{OR}\right]\right]=\sqrt{\frac{1}{r_T}+\frac{1}{n_T-r_T}+\frac{1}{r_C}+\frac{1}{n_C-r_C}}$$

This can be derived as follows:

$$Var\left[\log_e\left[\hat{OR}\right]\right]=Var\left[\log_e\left[\frac{p_T(1-p_C)}{(1-p_T)p_C}\right]\right]$$

$$=Var\left[\log_e\left[\frac{p_T}{(1-p_T)}\right]-\log_e\left[\frac{p_C}{(1-p_C)}\right]\right]$$

$$=Var\left[\log_e\left[\frac{p_T}{1-p_T}\right]\right]+Var\left[\log_e\left[\frac{p_C}{1-p_C}\right]\right](*)$$

because treatment groups are independent. Approximate standard errors can be calculated using the *Delta Method*, which is based on a Taylor Series approximation. This states that

$$Var\left[f(x)\right]\cong f'(x)^2_{x=E[x]}Var[x].$$

Considering $f(p_T)=\log_e\left[\frac{p_T}{1-p_T}\right]$,

Hence $f'(p_T)=\frac{1}{p_T}+\frac{1}{1-p_T}=\frac{1}{p_T(1-p_T)}$

Since $E[p_T]=\pi_T$ and $Var[p_T]=\frac{\pi_T(1-\pi_T)}{n_T}$,

it follows that

$$Var\left[\log_e\left[\frac{p_T}{1-p_T}\right]\right]=\left(\frac{1}{\pi_T(1-\pi_T)}\right)^2\frac{\pi_T(1-\pi_T)}{n_T}=\left(\frac{1}{n_T\pi_T(1-\pi_T)}\right).$$

Similarly,

$$Var\left[\log_e\left[\frac{p_C}{1-p_C}\right]\right]=\left(\frac{1}{n_C\pi_C(1-\pi_C)}\right).$$

Substitution in the equation (*) above give

$$Var\left[\log_e\left[\hat{OR}\right]\right]=\frac{1}{n_T\pi_T(1-\pi_T)}+\frac{1}{n_C\pi_C(1-\pi_C)}$$

$$=\frac{1}{n_T\pi_T}+\frac{1}{n_T(1-\pi_T)}+\frac{1}{n_C\pi_C}+\frac{1}{n_C(1-\pi_C)}.$$

The standard error can be obtained by substitution or p_T and p_C for π_T and π_C.

Hence $SE\left[\log_e\left[\hat{OR}\right]\right]=\sqrt{\frac{1}{n_Tp_T}+\frac{1}{n_T(1-p_T)}+\frac{1}{n_Cp_C}+\frac{1}{n_C(1-p_C)}}$

$=\sqrt{\frac{1}{r_T}+\frac{1}{n_T-r_T}+\frac{1}{r_C}+\frac{1}{n_C-r_C}}$ as required ■

Using this result the (1-α) confidence interval of $\log_e[OR]$ is

$$\log_e\left[\frac{r_T(n_C-r_C)}{(n_T-r_T)r_C}\right]\pm z_{\alpha/2}\sqrt{\frac{1}{r_T}+\frac{1}{n_T-r_T}+\frac{1}{r_C}+\frac{1}{n_C-r_C}}$$

Confidence intervals for the odds ratio are obtained by taking the exponents.

Ex 3.8 *For the data from the Propranolol trial calculate the 95% confidence interval for the odd of survival at 28 days for propranolol as compared to placebo.*

$\log_e\left[\hat{OR}\right] =$ $\log_e[3.1823] = 1.15759$

$\hat{SE}\left[\log_e\left[\hat{OR}\right]\right] = \sqrt{\frac{1}{r_T} + \frac{1}{n_T - r_T} + \frac{1}{r_C} + \frac{1}{n_C - r_C}} = \sqrt{\frac{1}{38} + \frac{1}{7} + \frac{1}{29} + \frac{1}{17}} = 0.5123$

95% Confidence intervals of $\log_e\left[\hat{OR}\right]$ are given by

$1.15759 \pm 1.96 \times 0.5123$, that is 0.1535 to 2.1617

Taking exponentials 95% Confidence interval of $\hat{OR}$ is (,)

(1.17 to 8.67)

Note we have calculated the odd ratio for survival. The odds ratio for death and its confidence interval are simple the reciprocal.

Hypotheses Test for the Odds Ratio

A test of the null hypothesis H_o:OR=1 could be based on the statistic $Z_{lor} = \frac{\log_e\left[\hat{OR}\right]}{\hat{SE}\left[\log_e\left[\hat{OR}\right]\right]}$ i.e. test $H_0 : \log_e\left[\hat{OR}\right] = 0$. In practice one does not do this as the test of the null H_o: OR=1 is equivalent to that the test based on the rate difference, H_o: RD=0, which is preferable as it does not depend an approximate standard error determined using the delta method. The p-value for the z-test of proportions calculated above for this data is therefore used for hypothesis tests of the odds ratio in preference to the statistic Z_{lor}.

Ex 3.9 Summarize the results of the propranolol trial based on Odds Ratios analysis.

" There was evidence that propranolol increased the odds of survival compared to placebo (OR= 3.18, 95% c.i. 1.17 to 8.67, ~~p=0.022).~~ 0.0208 from Z-test in Ex 3.5 on page 42

Interpretation of the Odds Ratio

Many people find odds ratios difficult to interpret, but the odds ratio is an important measure of effect in medical statistics. One reason for this is that the odds ratio can be estimated by logistic regression, which enables estimation of the odds ratio adjusted for other variables. It is very important for observational studies as it enables adjustment of effects for confounding variables. What is more the odds ratio is essential for case control studies as it is not possible to estimate either the risk difference or the risk ratio of the outcome due to the way in which subjects have been selected.

3.4 Analyses Based on the Rate Ratio

The rate ratio $RR = \frac{\pi_T}{\pi_C}$ can be estimated by $\hat{RR} = \frac{p_T}{p_C}$

Confidence intervals for the rate ratio, also called the risk ratio, can be constructed in a similar way to the odds ratio. As with the odds ratio the hypotheses test for the rate ratio is equivalent to that for the rate difference (*RD*) and so the z-test for proportions is still used.