

# NZ SRW Calving interval report

*Anthony Davidson*

*20 September, 2017*

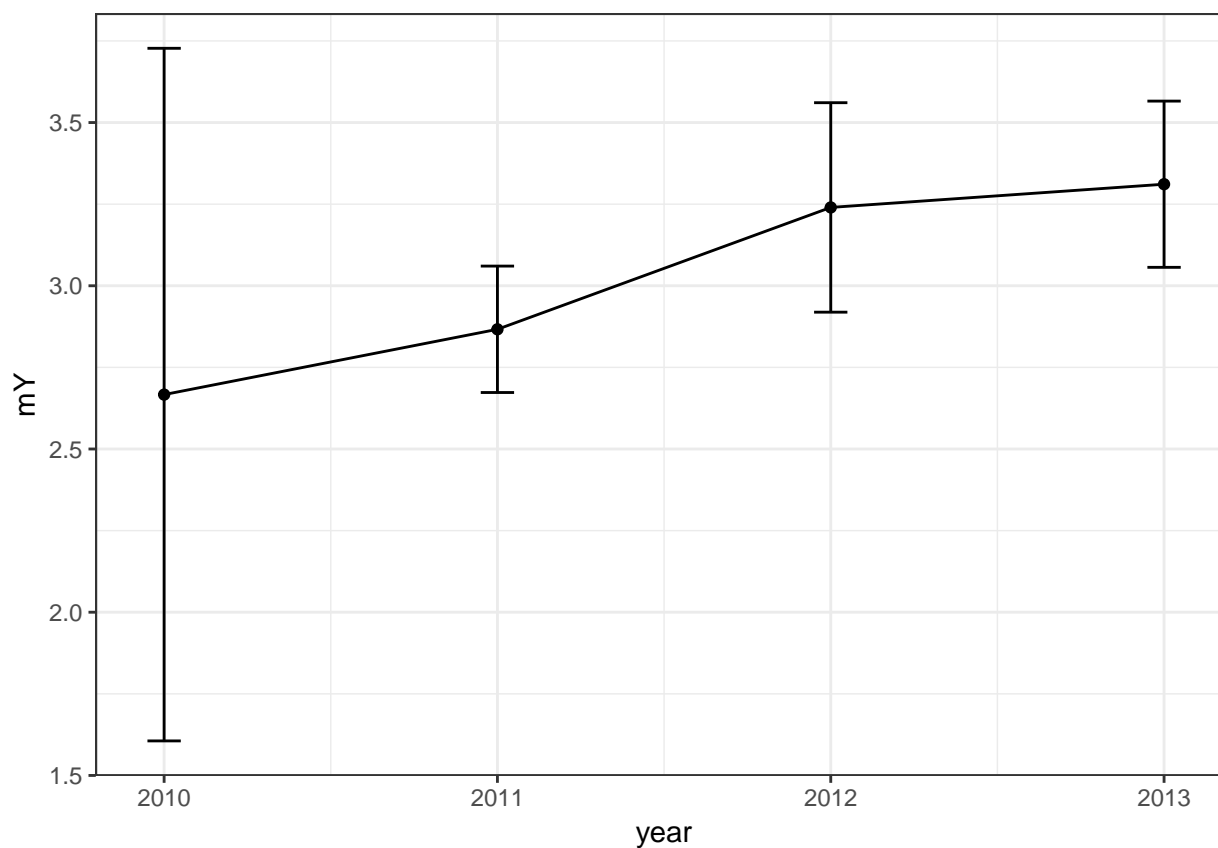
## Raw data

The data is then converted into intervals for each of the years 2010, 2011, 2012 and 2013.

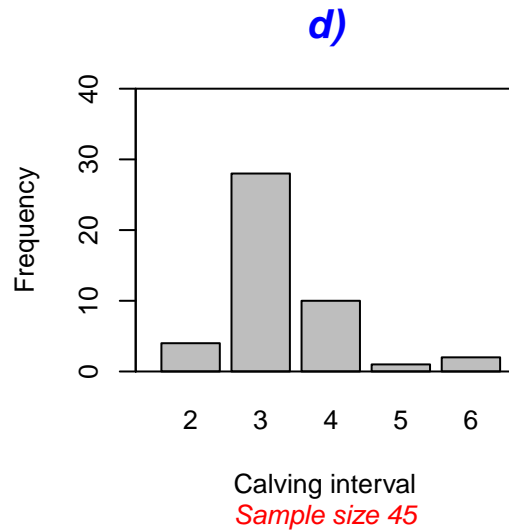
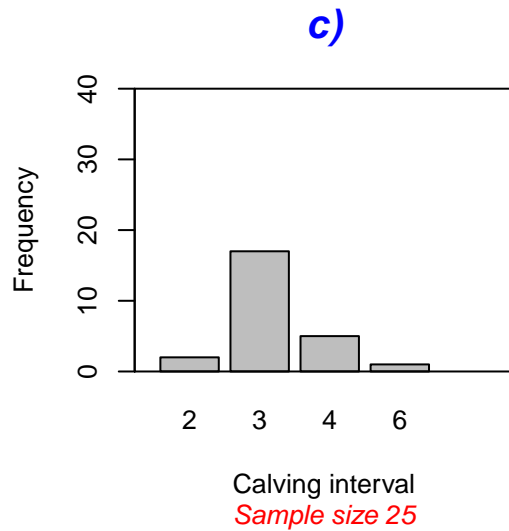
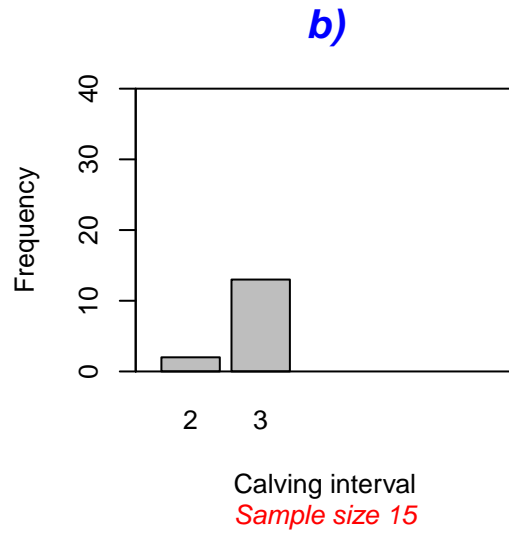
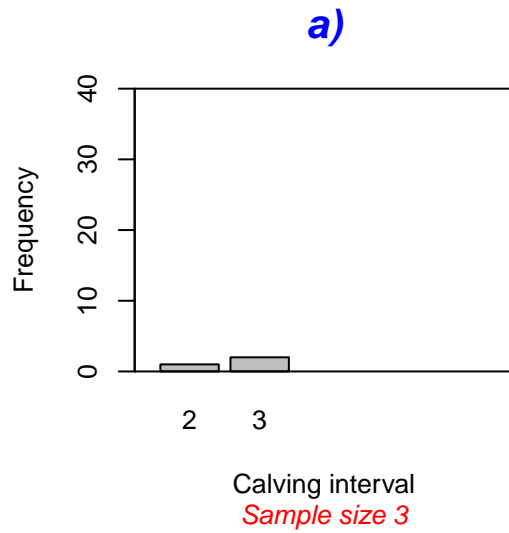
A table of the raw results as the year increase

year	n	mY	low.qt	high.qt	sd
2010	3	2.666667	1.605851	3.727482	0.5773503
2011	15	2.866667	2.673022	3.060312	0.3518658
2012	25	3.240000	2.919170	3.560830	0.7788881
2013	45	3.311111	3.056488	3.565734	0.8480518

The raw intervals can be presented as different means and confidence intervals as the length of the study increases:



We can also plot these as bar charts that show this more obviously:

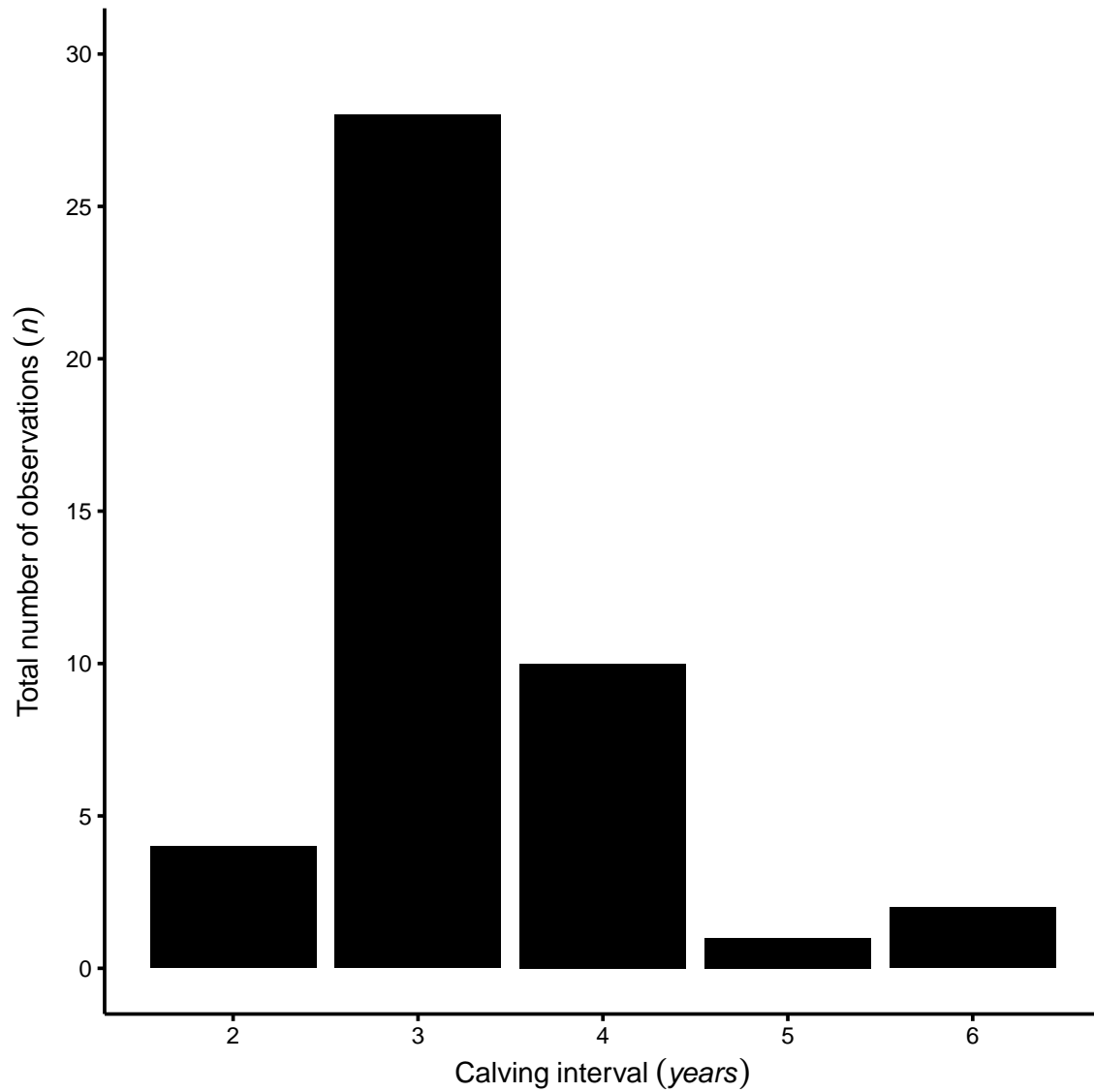


What is interesting is that in 2011 there were more intervals that subsequently reduced the standard error of the estimate and the precision increased.

**My idea of why this has happened is as follows:**

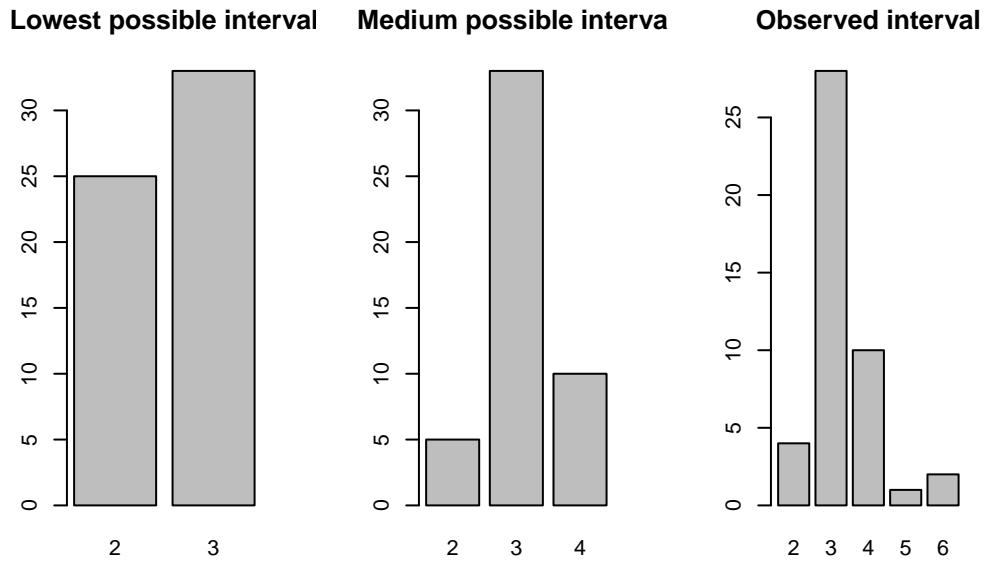
At the time of collection in 2011 the number of possible intervals greater than 5 was very unlikely (as there were only 5 years of research). This meant that the overall error of the estimate was reduced as there was very few “chances” of obtaining a calving interval of 5 or 6 even though there may have been quite a few. The following estimates then picked these up and the error in the estimate increased again.

Publication plot (Figure 2)



### Missing intervals (Bradford et al. 2008)

This is a way to see what the calving interval might be if we had in fact missed calving events that happened before or after the study period.



Publication plot (Not included in final publication)



The error gets smaller as the variance in the estimate decreases. We are artificially doing this here by modifying data to match what maybe more biologically possible.

Interval	Sample size	Mean	Lower limit	Higher limit	SD
Low	58	2.568966	2.437666	2.700265	0.4995461
Medium	48	3.104167	2.943089	3.265244	0.5550382
Observed	45	3.311111	3.056488	3.565734	0.8480518

## Bootstrapping

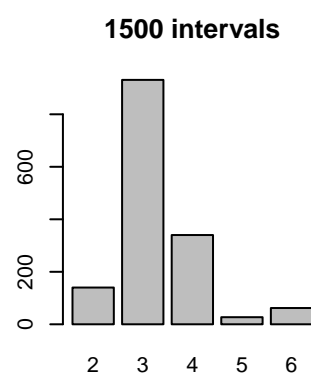
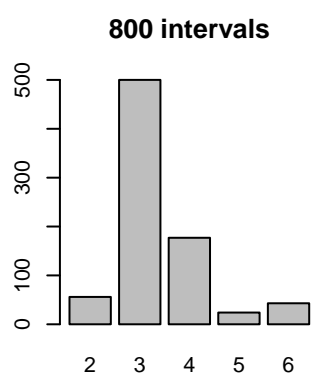
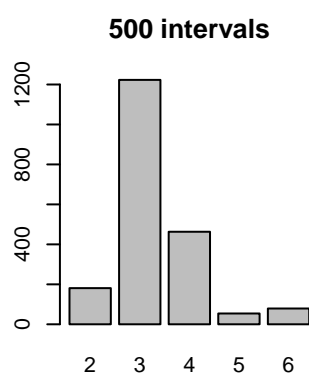
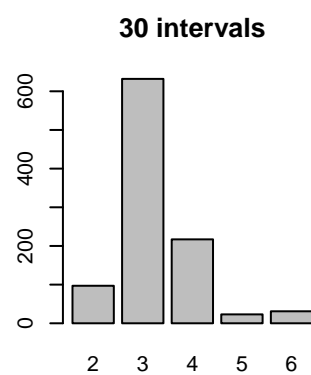
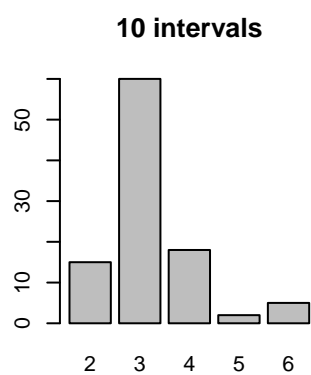
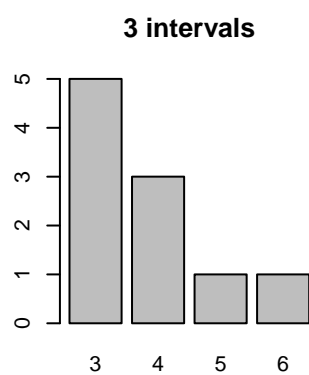
Below is the estimates from other populations of SRWs around the world.

Sample size	Mean	Lower limit	Higher limit	SE	Author	Location
NA	3.12	3.07	3.17	NA	Best et al. 2001	South Africa
1504	3.15	3.11	3.18	NA	Best et al. 2005	South Africa (1971-2003 Updated)
NA	3.16	3.13	3.19	NA	Brandao et al 2010	South Africa ( 1971-2006 Updated)
NA	3.35	NA	NA	0.05	Cooke et al. 2001	Argentina
749	3.42	NA	NA	0.11	Cooke et al. 2003	Argentina
NA	3.63	NA	NA	0.13	Burnell 2001	Australia

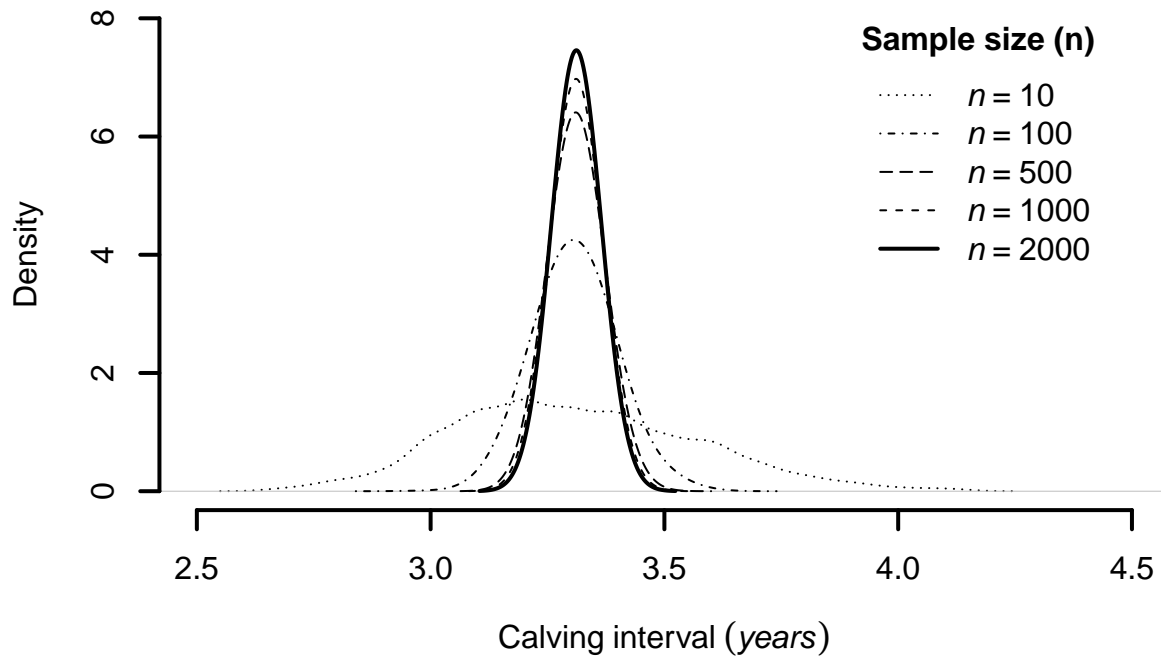
**Table of parameters:** This is taken directly from my Master's thesis. There may be updated estimates that need to be checked here.

## Single bootstrap with increased sample size

Bootstrap mean calving interval 1000 times and save the mean for each bootstrap sample. Here I investigate the effect of a sample size of 10,100,1000,2000 from the observed NZ calving interval.



Publication plot (Figure 3)



## Referees additional analysis

### Suggestion one

“If another 45 samples were collected over the next 8 years, but the mean moved from 3.31 to e.g., 3.7 years, would it be detectable with  $\alpha = 0.05, 0.10 \dots$  ? Given that other authors have hypothesised that poor food availability would translate to increased calving intervals/calf mortalities (e.g., Cooke et al 2003, Rowntree et al 2013), being able to detect such a change would be important.”

To do this it is simply a old school power analysis with the focus on the type one error ( $\alpha$ ) as below:

```
#observed sample
rev.one <- bdata$RealCI[1:45]

#sample 45 times
sample.true <- year2013$interval

#power analysis
pwr.test.results <- power.t.test(n = 45, # sample size
  delta = seq(0,0.99,0.001), #difference between means
  sd = sd(sample.true),      #observed variation
  alternative = "one.sided", #observed test type
```

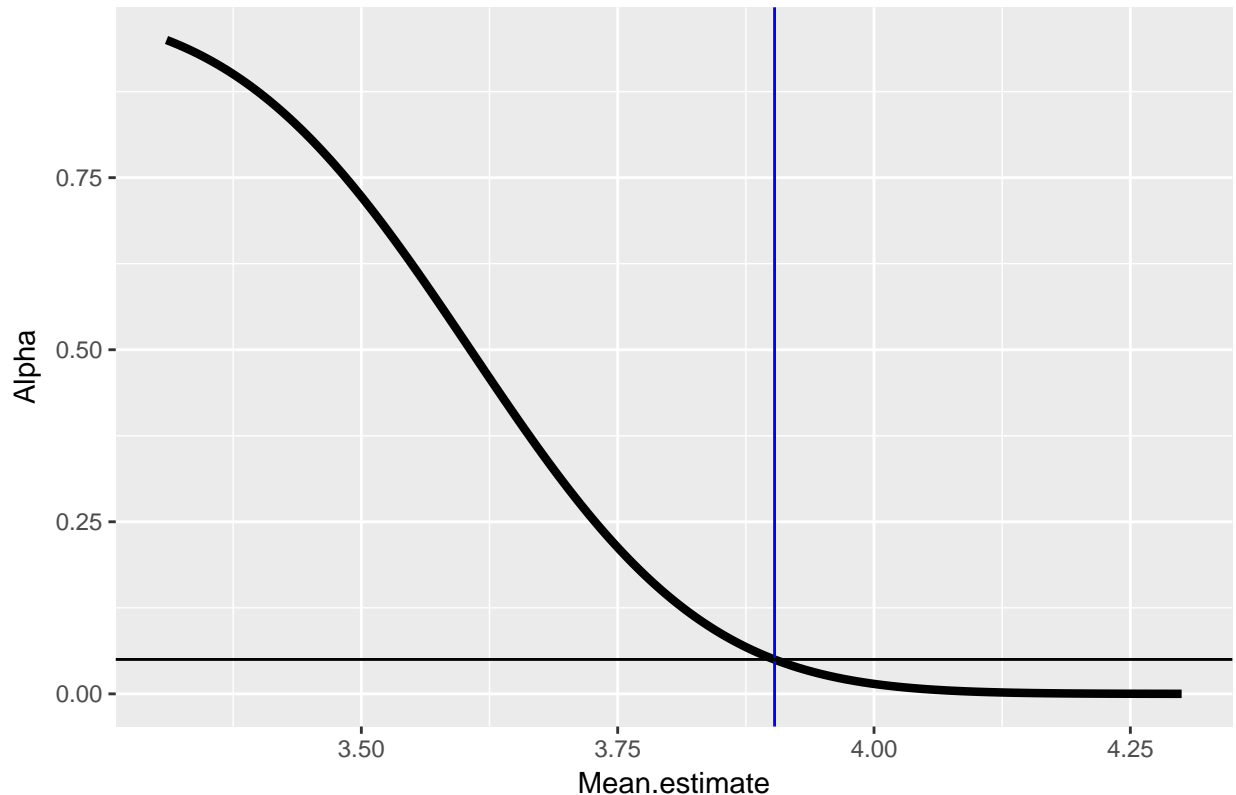
```
sig.level = 0.05)           #significance level
```

*#additional packages are available for more complex analysis  
#but have not done this as don't think it is needed*

Here I have looked at the effect of the difference in the mean to range from the same 3.31 (delta = 0) to 4.31 (delta = 1) with the same variation as observed in the observed data.

```
##      Power Mean.difference      Alpha Mean.estimate
## 1 0.9501505          0.593 0.04984946          3.903
```

Raw data result plot (n = 45)

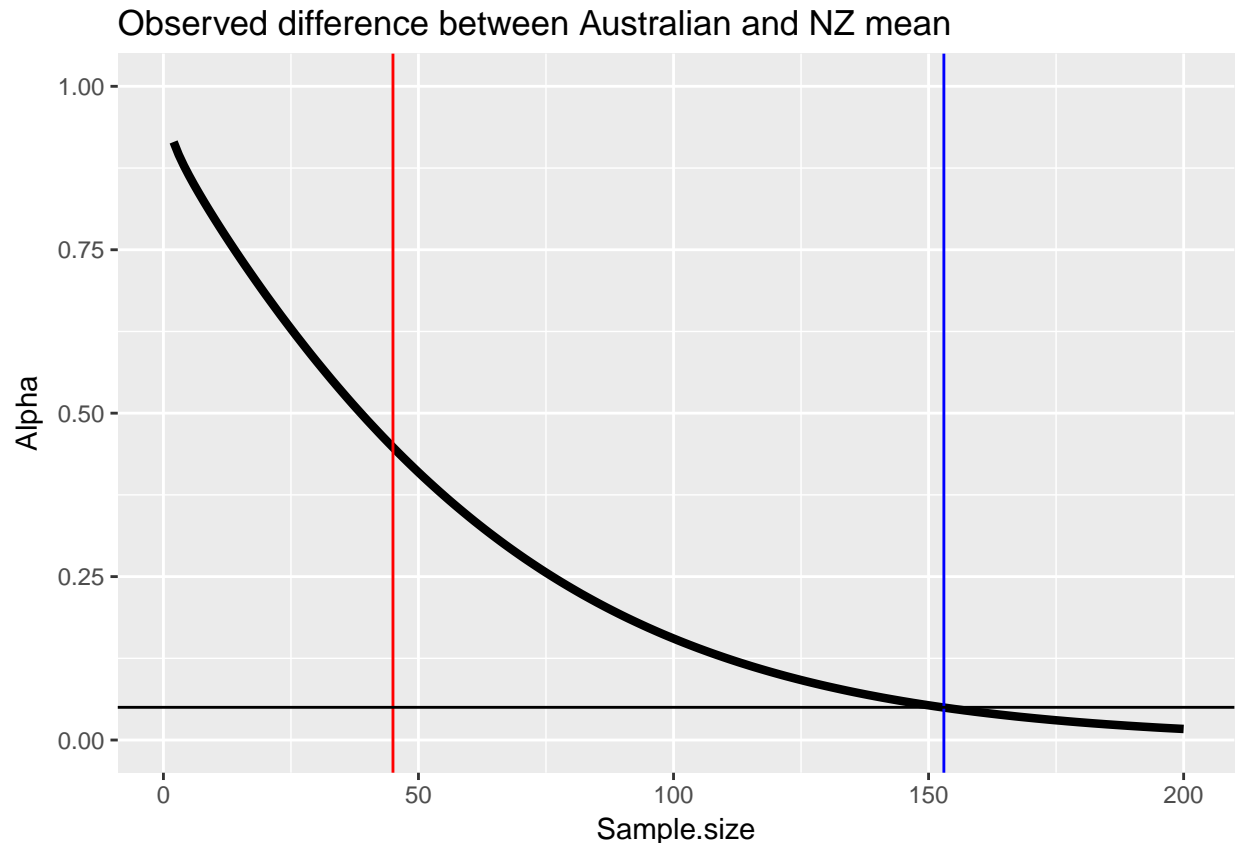


It is interesting to note that with the minimum mean difference with a sample size of 45 that a difference between populations was observed at the 95% level would be a mean calving interval of 3.9.

Below is an additional analysis to look at the number of samples needed to find a difference the same or greater than the mean estimate for the Australian population recorded as 3.63.

```
##      Power Sample.size      Alpha
## 1 0.9503366          153 0.0496634
```





Red line represents the current sample size of NZ population estimate and blue line represents the size needed to have  $\alpha=0.05$

### Important note

I do not this is the best approach because...

if the sampling continues over the next 8 years, the number of observations will not increase in a linear function as the number of individuals available for recapture is growing (more individuals in database) at a different rate to the capture probability (that is hopefully constant). I will do the simulation recommended because both the editor and reviewer wanted it.

I previously approached this problem by using bootstrapping and looking at the effect of the sampling size (n) increasing.

### Suggestion three

**“Is there any information for the females with the potentially missed calving events in the intervening years; e.g., were they seen without calves?”**

### Individual status

I only have data for these overall data points up to 2012 so have done the analysis with this and can just add an additional year to check this.

To do this the following steps are taken:

- 1) Raw data I have currently looks like so:

```
## Observations: 180
## Variables: 10
## $ ID      <fctr> AI06006, AI06007, AI06015, AI06022, AI06038, AI10...
## $ X2006    <int> 1, 2, 2, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ X2007    <int> 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ X2008    <int> 1, 0, 1, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0,...
## $ X2009    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ X2010    <int> 0, 0, 2, 0, 0, 6, 6, 5, 5, 3, 4, 2, 5, 4, 5, 3, 2,...
## $ X2011    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ X2012    <int> 0, 1, 2, 4, 0, 0, 0, 0, 0, 0, 5, 0, 0, 12, 0, 0, 0...
## $ total    <int> 2, 4, 8, 5, 3, 6, 7, 5, 5, 3, 9, 2, 5, 16, 6, 3, 2...
## $ X..yrs.seen <int> 2, 3, 5, 2, 2, 1, 2, 1, 1, 1, 2, 1, 1, 2, 2, 1, 1,...

## Observations: 41
## Variables: 8
## $ ID      <fctr> AI10124, AI10070, AI10086, AI08340, AI08341, AI...
## $ Yr.first.seen <int> 2007, 2008, 2007, 2008, 2008, 2008, 2008, 2008, ...
## $ Calves    <int> 2007, 2008, 2007, 2008, 2008, 2008, 2008, 2008, ...
## $ Calves.1  <int> 2010, 2010, 2010, 2011, 2011, 2011, 2011, 2011, ...
## $ Calves.2  <int> 2013, 2013, 2013, NA, NA, NA, NA, NA, NA, NA, NA...
## $ Interval.1 <int> 3, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 6, ...
## $ Interval.2 <int> 3, 3, 3, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ X        <lg1> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

- 2) The two data sets can be combined to reduce the data down to just individuals that have been seen with a calf at least once.

This is important because we do not know the reproductive state of individuals until they are first seen with a calf so therefore any individuals that are in the database and have not been identified with a calf maybe juvenile and therefore not reproducing and we would be overestimating the number of possible reproductive events.

```
## Observations: 237
## Variables: 7
## $ ID      <chr> "AI06006", "AI06007", "AI06015", "AI06022", "AI0...
## $ year     <chr> "X2006", "X2006", "X2006", "X2006", "X2006", "X2...
## $ count    <int> 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ Yr.first.seen <int> NA, NA, NA, 2006, NA, NA, NA, 2007, 2007, NA, NA...
## $ Calves    <int> NA, NA, NA, 2006, NA, NA, NA, 2007, 2007, NA, NA...
## $ Calves.1  <int> NA, NA, NA, 2012, NA, NA, NA, 2013, 2010, NA, NA...
## $ Calves.2  <int> NA, NA, NA, NA, NA, NA, NA, NA, 2013, NA, NA, 20...

## # A tibble: 1 × 4
##       ID year Calves Calves.1
##   <chr> <chr> <int>    <int>
## 1 AI09216 2007  2009    2011
```

- 3) And overall between 2006 and 2012 only a single individual (AI09216) was identified as being a mother and was seen in 2007 without a calf.

This individual had also not been seen with a calf until 2009 so even this individual may have been juvenile at this stage.

## Suggestions on variations of detection error and removing/increasing observed estimates

These comments are about additional data manipulation and the following estimates are plotted in the same figure as the Bradford et al 2008 modifications below.

“Another possibility not explored is that the reported two year calving intervals could have been an error; what would the mean calving interval be in this case?”

```
greater.than.2 <- sample.true[sample.true>2]

#greater.than.2
mean.2<-sum(greater.than.2)/length(greater.than.2)
s.2<-sd(greater.than.2)
SE.2<-s2013/(sqrt(length(greater.than.2)))
n.2<-length(greater.than.2)
low.qt.2<- mean.2-(qt(0.975,length(greater.than.2))*SE.2)
high.qt.2 <- mean.2+(qt(0.975,length(greater.than.2))*SE.2)

#add it to the table from bradford data
Sumtable[4,] <- c("miss2year",n.2,mean.2,low.qt.2,
                  high.qt.2,sd(greater.than.2))
```

“The existing simulation could be improved by assuming imperfect detection of calving events. For example, how does the variance around the estimate of mean calving interval change when 5% of calving events are not detected, increasing the number of 5 or 6 year observed calving intervals?”

```
##### 2.2%
#parameters
boots <- 1000
n <- c(1:1000)

###round all percentages upwards
detect1 <- 44 # (45*1.02) - 45 = 0.9
detect2 <- 42 # (45*1.05) - 45 = 2.25
detect3 <- 40 # (45*1.10) - 45 = 4.5

sample2 <-rep(NA, 1000)
sample5 <-rep(NA, 1000)
sample10 <-rep(NA, 1000)

for (i in 1:boots) {
  sample2[i]<-mean(sample(year2013$interval,detect1,replace=T))
  sample5[i]<-mean(sample(year2013$interval,detect2,replace=T))
  sample10[i]<-mean(sample(year2013$interval,detect3,replace=T))
} #i

#####estimates#####
sample2 <- sort(sample2)
#low = 25/1000
sample2.2.5 <- sample2[25]
#median
sample2.50 <- sample2[500]
#high = 975/1000
```

```

sample2.975 <- sample2[975]

sample5 <- sort(sample5)
#low = 25/1000
sample5.2.5 <- sample5[25]
#median
sample5.50 <- sample5[500]
#high = 975/1000
sample5.975 <- sample5[975]

sample10 <- sort(sample10)
#low = 25/1000
sample10.2.5 <- sample10[25]
#median
sample10.50 <- sample10[500]
#high = 975/1000
sample10.975 <- sample10[975]

#add it to the table from bradford data
Sumtable[5,] <- c("detect1",detect1,sample2.50,sample2.2.5,sample2.975,NA)
Sumtable[6,] <- c("detect2",detect2,sample5.50,sample5.2.5,sample5.975,NA)
Sumtable[7,] <- c("detect5",detect3,sample10.50,sample10.2.5,sample10.975,NA)

```

“or increasing the number of 5 or 6 year observed calving intervals?”

Here I have added twice as many 5 and 6 year intervals (5yr = 2, 6yr = 4)

```

longer5.6 <- c(sample.true,5,6,6)

#greater.than.2
mean.56<-sum(longer5.6)/length(longer5.6)
s.56<-sd(longer5.6)
SE.56<-s.56/(sqrt(length(longer5.6)))
n.56<-(length(longer5.6))
low.qt.56<- mean.56-(qt(0.975,length(longer5.6))*SE.56)
high.qt.56 <- mean.56+(qt(0.975,length(longer5.6))*SE.56)

#add it to the table from bradford data
Sumtable[8,] <- c("longer.56",n.56,mean.56,low.qt.56,high.qt.56,sd(longer5.6))

###sort out numbering in dataframe
Sumtable <- as.data.frame(Sumtable)

Sumtable$n <- as.numeric(as.character(Sumtable$n))
Sumtable$mY <- as.numeric(as.character(Sumtable$mY))
Sumtable$low.qt <- as.numeric(as.character(Sumtable$low.qt))
Sumtable$high.qt <- as.numeric(as.character(Sumtable$high.qt))
Sumtable$sd <- as.numeric(as.character(Sumtable$sd))
Sumtable$interval <- as.character(Sumtable$interval)

```

Table of results

Interval	Sample size	Mean	Lower limit	Higher limit	SD
Low	58	2.568966	2.437666	2.700265	0.4995461
Medium	48	3.104167	2.943089	3.265244	0.5550382
Observed	45	3.311111	3.056488	3.565734	0.8480518
miss2year	41	3.439024	3.171549	3.706499	0.7761695
detect1	44	3.295454	3.090909	3.568182	NA
detect2	42	3.309524	3.071429	3.571429	NA
detect5	40	3.325000	3.050000	3.600000	NA
longer.56	48	3.458333	3.165307	3.751360	1.0097047

Publication plot (Not included in final publication)

