

# L9: Resampling methods



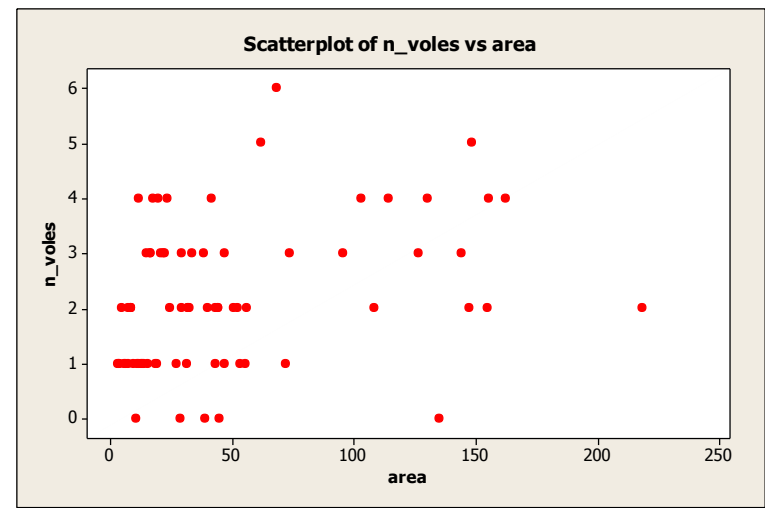
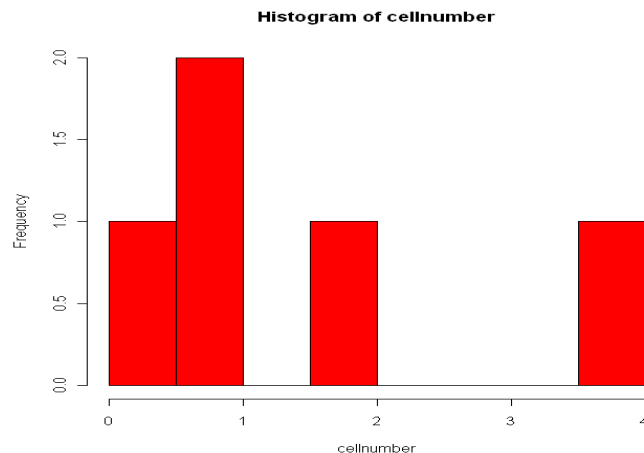
# Random Resampling

**Computationally-intensive** statistical techniques that avoid parametric assumptions by extracting maximum information from the data:

- We *re-sample* our original data many times and calculate a statistic of interest (mean, variance, correlation coefficient, etc.).
- Each re-sample contains a random component and the many re-samples give a new distribution



- When **sample sizes are small** or the underlying **distributions are non-normal** or contain **outliers**, standard parametric techniques can generate misleading results.
- **Resampling gives us robust and powerful** techniques to avoid these problems. We can:
  - **reduce uncertainty in descriptive statistics** (means, variances, confidence intervals, etc)
  - **do tests** ( $t$ ,  $r$ ,  $F$ , etc) **against a “perfect” null distribution**
- Resampling uses **ONLY** the information in your data, not assuming anything about the distribution they came from.



Resampling to reduce  
uncertainty in  
descriptive statistics

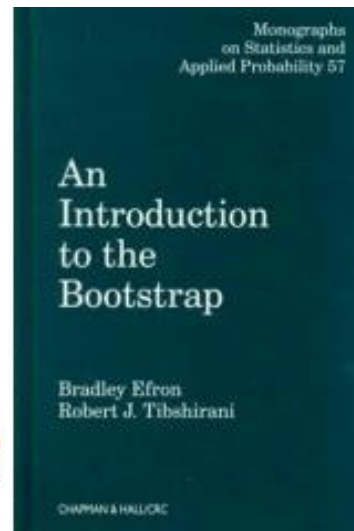
# Bootstrap and Jack-knife

- Bootstrapping is for estimating confidence intervals around parameter estimates
- Jack-knifing is for sensitivity analysis: identify outliers and their impact

Method	Description	Nickname
Bootstrap	Randomly resample with replacement	“sample with replacement”
Jackknife	Systematically omit single data points	“drop one out”

# The Bootstrap

- For estimating the **variation** in a test statistic (e.g. SE and 95% CI) without assuming a distribution
- Resample the observed data **with replacement** (*aka Monte Carlo sampling*) a large number ( $\geq 1000$ ) of times, making each the same size as the original sample
- Calculate the statistic of interest for each of the samples, yielding a **bootstrapped distribution** for the statistic.



# Grouse on leks

Imagine we have these data  
for numbers of black  
grouse on each of 20 leks:

Number <-

```
c(0,1,3,1,1,0,0,8,10,4,1,5,  
0,8,0,1,1,0,1,5)
```

- mean and standard error of this distribution is  $2.500 \pm 0.698$
- Para 95% CI =  $SE \cdot t_{crit}$   
 $= 0.698 \times 2.093 = 1.46$   
 $\rightarrow \{1.03-3.96\}$

DF	P=0.1	P=0.05	P=0.01	P=0.001
1	6.314	12.706	63.657	636.619
2	2.92	4.303	9.925	31.599
3	2.353	3.182	5.841	12.924
19	1.729	2.093	2.861	3.883
20	1.725	2.086	2.845	3.85
infinity	1.645	1.96	2.576	3.291

How robust are these estimates?

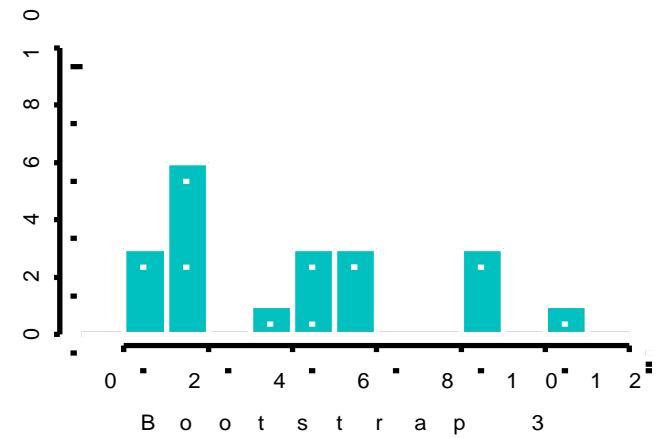
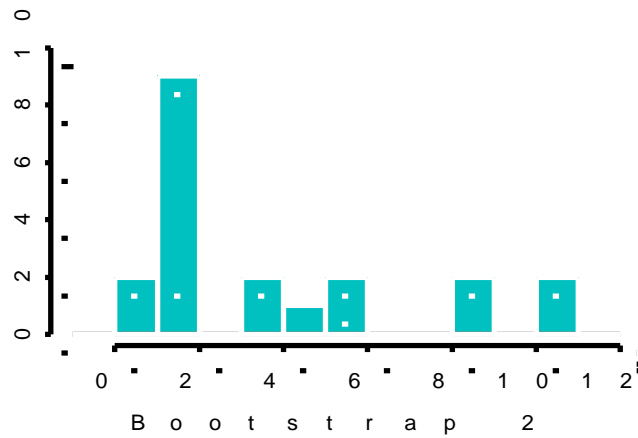
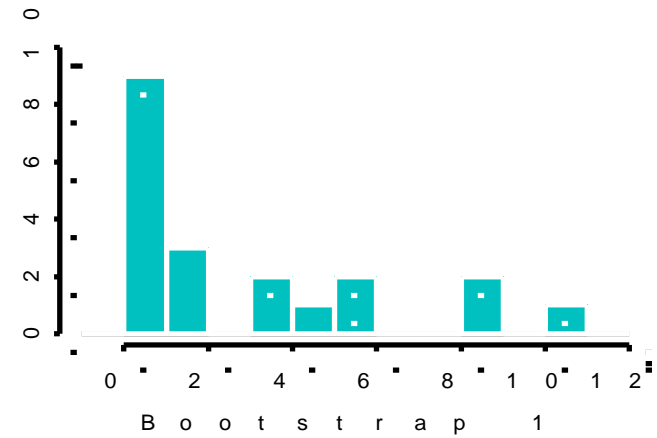
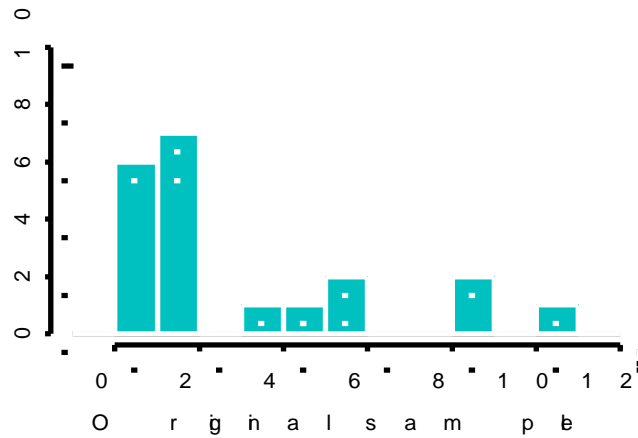


# Sampling with replacement

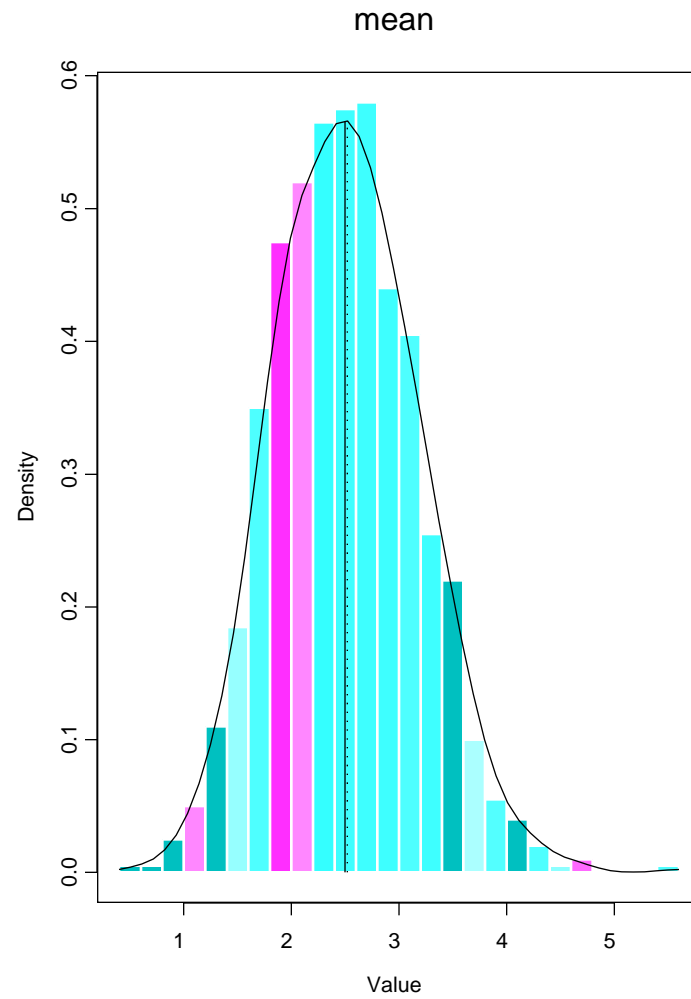
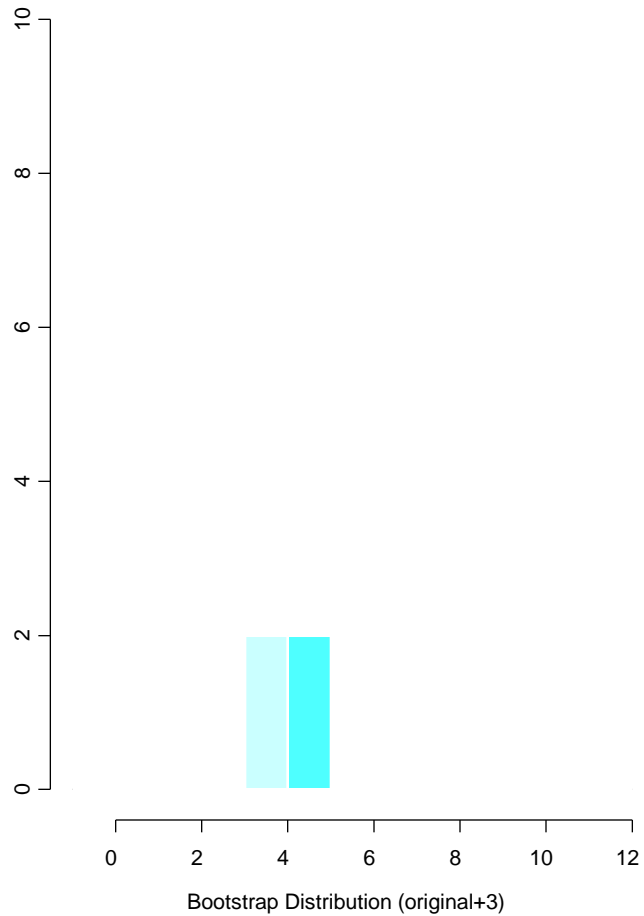
<i>Lek No.</i>	<i>Lek Size</i>	Bootstrap 1	Bootstrap 2	Bootstrap 3
1	0	0	0	1
2	1	3	3	5
3	3	0	1	8
4	1	10	8	1
5	1	5	0	4
6	0	0	10	1
7	0	8	5	0
8	8	0	1	1
9	10	8	5	5
10	4	1	8	5
11	1	3	4	0
12	5	0	1	3
13	0	0	3	10
14	8	1	1	8
15	0	0	1	4
16	1	1	1	8
17	1	0	10	1
18	0	0	1	0
19	1	5	1	1
20	5	4	1	4
<b>Mean</b>	<b>2.500</b>	<b>2.450</b>	<b>3.25</b>	<b>3.500</b>
<b>standard error</b>	<b>0.698</b>	<b>0.716</b>	<b>0.743</b>	<b>0.698</b>



# Bootstrap histograms



# Bootstrap distribution



```
> b <- boot(Number, statistic=samplemean, R=1000)
> b
```

ORDINARY NONPARAMETRIC BOOTSTRAP

```
Call:
boot(data = Number, statistic = samplemean, R = 1000)
```

```
Bootstrap Statistics :
      original    bias      std. error
t1*          2.5 -0.0083     0.6683733
```

```
> boot.ci(b, conf=0.95, type = c("norm", "basic", "perc", "bca"))
```

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS

Based on 1000 bootstrap replicates

CALL :

```
boot.ci(boot.out = b, conf = 0.95, type = c("norm", "basic",
      "perc", "bca"))
```

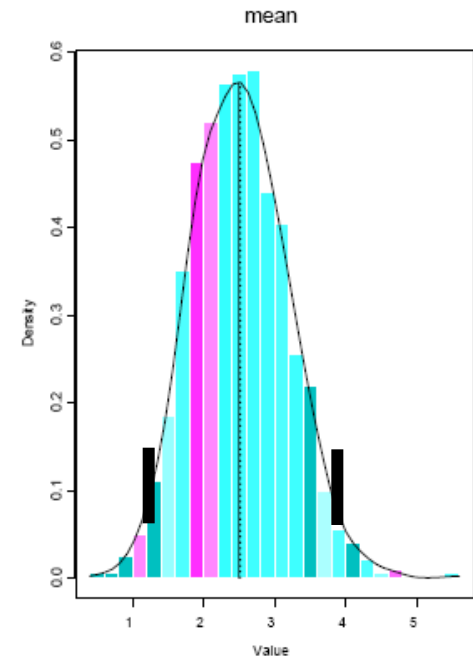
Intervals :

Level	Normal	Basic
95%	( 1.198, 3.818 )	( 1.100, 3.749 )

Level	Percentile	BCa
95%	( 1.251, 3.900 )	( 1.400, 4.147 )

Calculations and Intervals on Original Scale

```
> samplemean <- function(x, d) {
+   return(mean(x[d]))
+ }
```

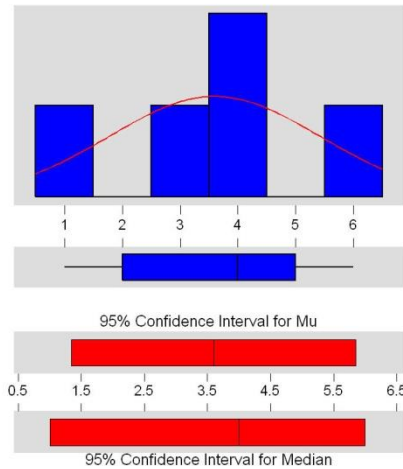


Mean= 2.5  
(95% CI 1.4-4.147)

Parametric estimates  
**95% CI 1.03-3.96**

# Descriptive Statistics

Variable: rand



Anderson-Darling Normality Test

A-Squared: 0.245  
P-Value: 0.570

Mean 3.60000  
StDev 1.81659  
Variance 3.3  
Skewness -2.7E-01  
Kurtosis 1.07438  
N 5

Minimum 1.00000  
1st Quartile 2.00000  
Median 4.00000  
3rd Quartile 5.00000  
Maximum 6.00000

95% Confidence Interval for Mu  
1.34441 5.85559

95% Confidence Interval for Sigma  
1.08838 5.22007

95% Confidence Interval for Median  
1.00000 6.00000

# A sample of just {4,6,4,1,3}...

```
> b <- boot(rand, samplemean, R=1000)
```

```
> b
```

ORDINARY NONPARAMETRIC BOOTSTRAP

Call:

```
boot(data = rand, statistic = samplemean, R = 1000)
```

Bootstrap Statistics :

	original	bias	std. error
t1*	3.6	0.025	0.7241681

```
> boot.ci(b, conf=0.95, type =  
+ c("norm", "basic", "perc", "bca"))
```

Types of  
confidence  
intervals...

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS  
Based on 1000 bootstrap replicates

CALL :

```
boot.ci(boot.out = b, conf = 0.95, type = c("norm",  
"basic",  
"perc", "bca"))
```

Intervals :

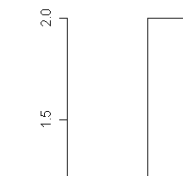
Level	Normal	Basic
95%	( 2.156, 4.994 )	( 2.200, 5.000 )

Level	Percentile	BCa
95%	( 2.2, 5.0 )	( 1.8, 4.6 )

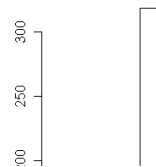
Calculations and Intervals on Original Scale

Some BCa intervals may be unstable

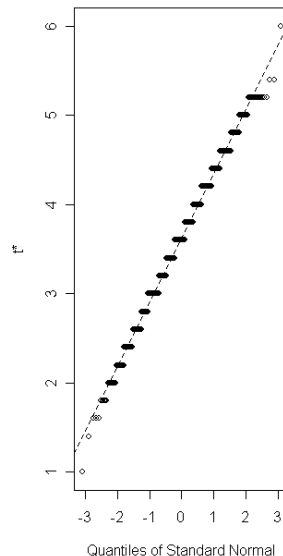
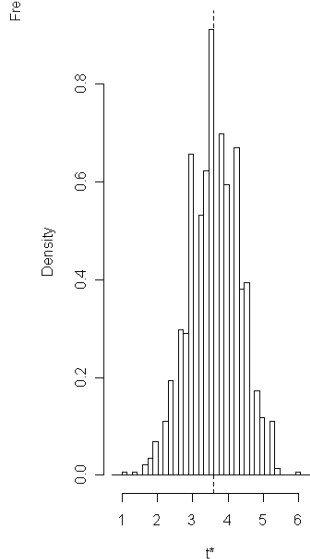
Histogram of rand



Histogram of b\$t



Histogram of t

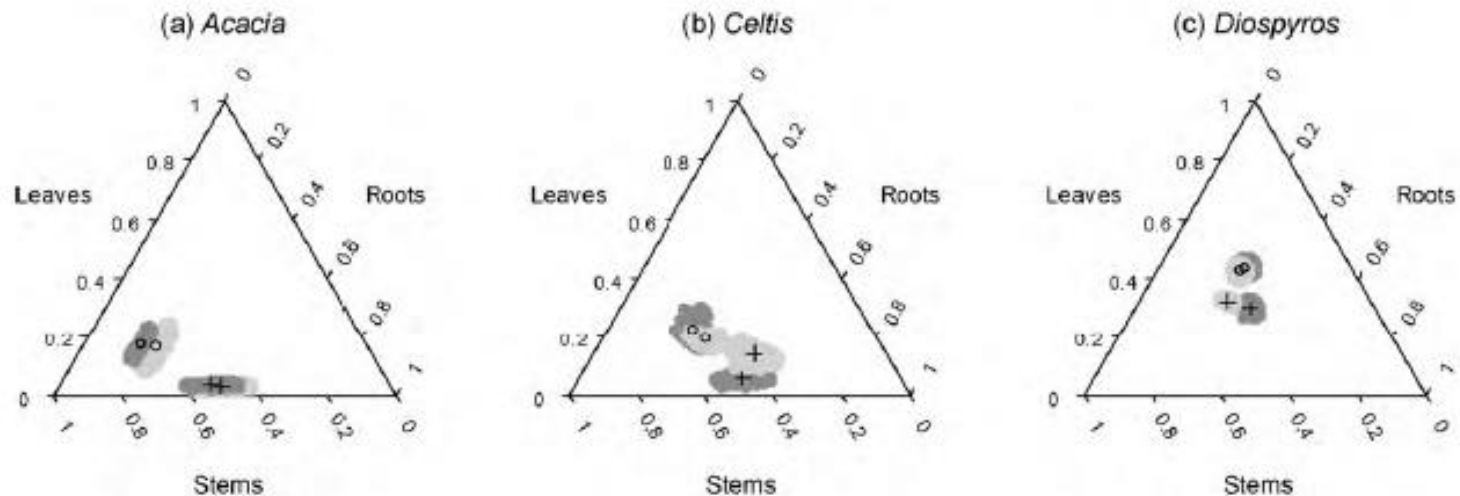


# Can estimate anything...

1. mean and variance of a single sample
2. difference between two samples
3. correlation between two samples
4. regression coefficients
5. ratio of variances, etc,

6

*R.M. Gunton et al. / Forest Ecology and Management xxx (2010) xxx–xxx*



**Fig. 3.** Mean leaf, root and stem mass fractions for the three species (shaded regions indicate 95% confidence regions from bootstrapping). The mass-partitioning patterns for each species are shown for low light (circles) and high light (crosses) with low nitrogen (pale shading) and high nitrogen (dark shading).

# Can estimate anything...

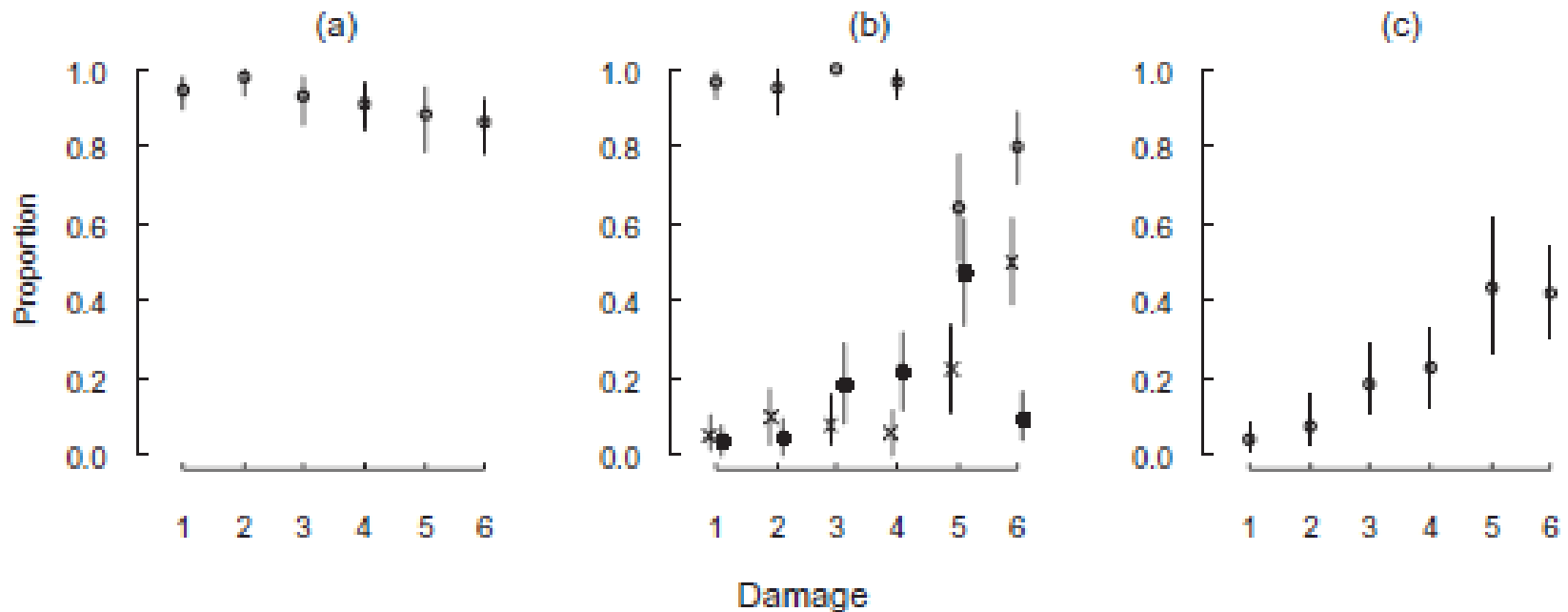


Fig. 3. Detailed results for *Eucalyptus tetrodonta* from the extended dataset: proportions of (a) trees that resprouted (b) trees with each type of resprout, among those that resprouted (root: crosses; basal: filled circles; epicormic: open circles) and (c) trees with root or basal resprouts, among those that resprouted epicormically. Codes for damage types: 1 = tertiary or higher-level branches survived, 2 = secondary branches survived, 3 = only primary branches survived, 4 = main stem snapped above 2 m from the ground, 5 = main stem snapped up to 2 m above ground level, 6 = uprooted. The vertical bars show 95% confidence intervals estimated from bootstrapping.

Franklin et al (2009), Austral Ecology

# The Jack-knife

How sensitive is a statistic to any one datum?

- useful for **identifying outliers** and deriving robust confidence intervals after compensating for the bias these outliers introduce.
- we recalculate the statistic of interest (mean,  $t$ ,  $r$ , etc)  **$N$**  times, where  $n$  is the number of data points we have. Each time omit one data point
- Thus our  **$N$  estimates are calculated using  $N-1$  observations each.**
- From our  $n$  estimates we can calculate a mean and an estimate of the variation in our data. Just as importantly, we can identify unusual (outlier) data points.



# Leaving data out

<i>Lek No.</i>	<i>Number of birds</i>	Jackknife 1	Jackknife 2	Jackknife 3
1	0	X	0	0
2	1	1	X	1
3	3	3	3	X
4	1	1	1	1
5	1	1	1	1
6	0	0	0	0
7	0	0	0	0
8	8	8	8	8
9	10	10	10	10
10	4	4	4	4
11	1	1	1	1
12	5	5	5	5
13	0	0	0	0
14	8	8	8	8
15	0	0	0	0
16	1	1	1	1
17	1	1	1	1
18	0	0	0	0
19	1	1	1	1
20	5	5	5	5
Mean	2.500	2.632	2.579	2.474
standard error	0.698	0.722	0.731	0.735

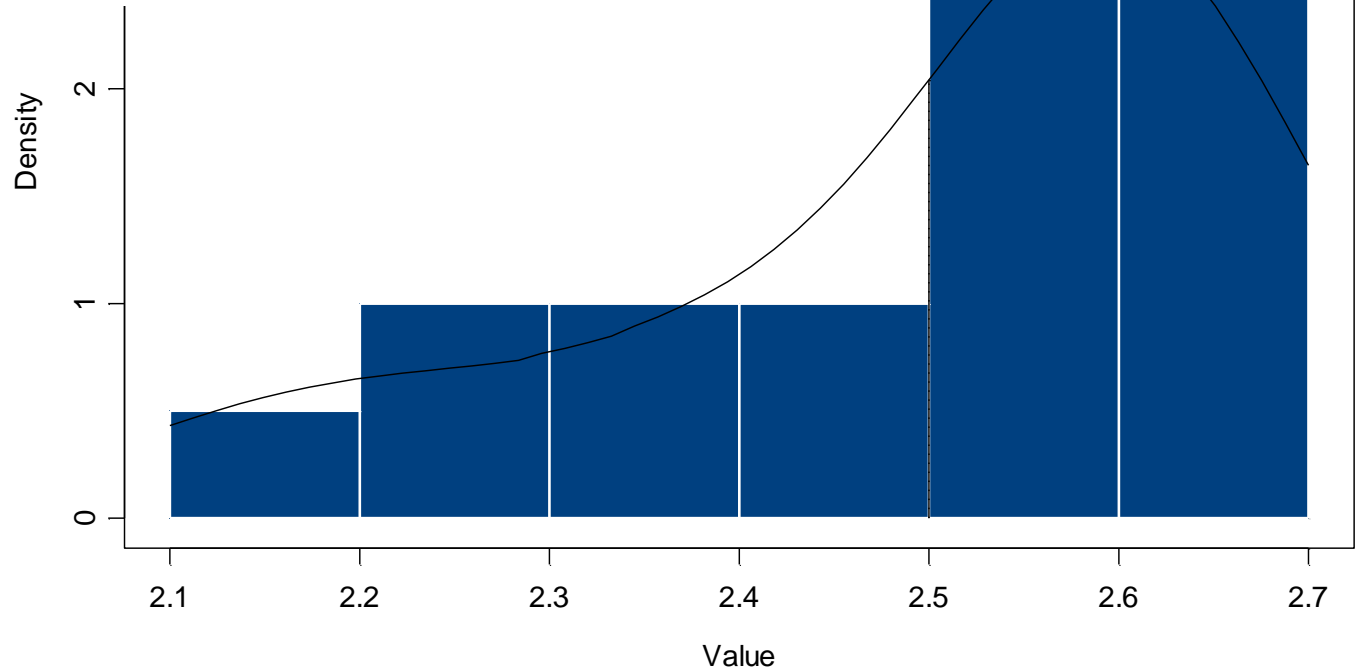


# Jack-knife distribution

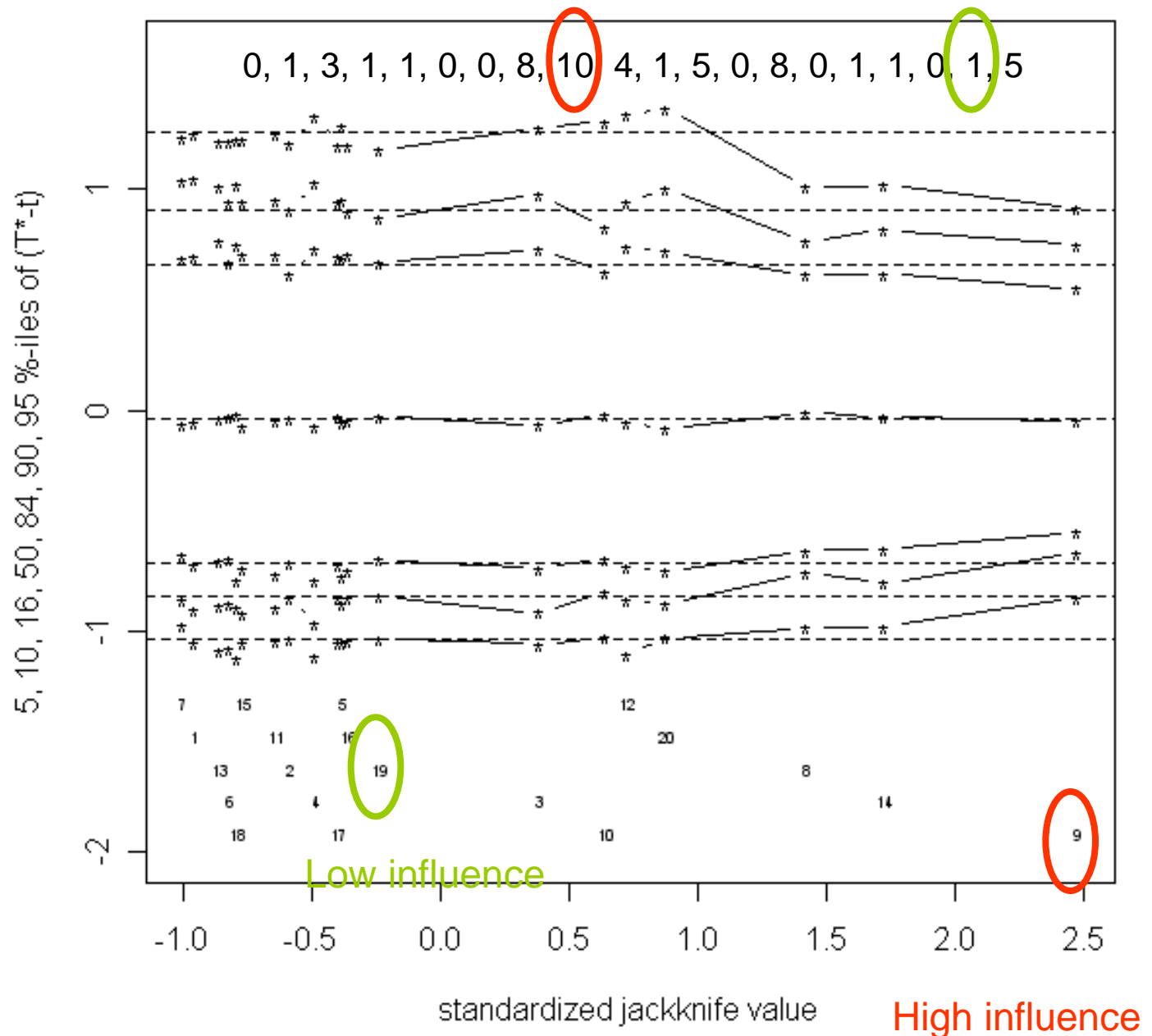
Observed mean = 2.5

**Empirical Percentiles:**

	2.5%	5%	95%	97.5%
mean	2.155	2.205	2.632	2.632



# jack.after.bootstrap()

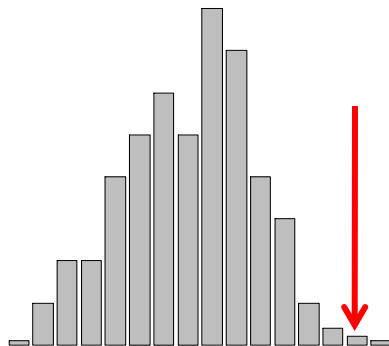


Randomisation:  
the “perfect null”

# Randomisation for the perfect null

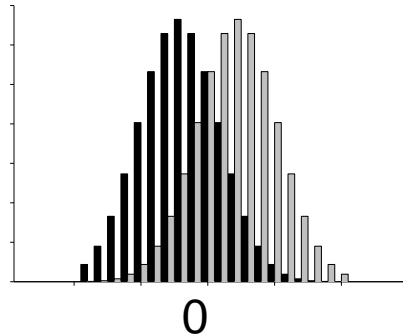
Implement your null hypothesis: build up a “null distribution” from your data to compare an observed statistic to.

- Allows you to get away from parametric assumptions and avoid low power of non-parametric tests
- Very flexible – customise your tests



# What is a null hypothesis?

- “A precise but inaccurate dream...”
- For Normal errors, we can take the observed  $s$  to estimate  $\sigma$  and then imagine  $\mu$  is 0.



- For GLMs, we can't do this since mean is linked to SD. So our null model already has an intercept, but we can use deviance estimates to imagine what the stats would be if all other effects are 0.
- **What about weird distributions?**

# DIY “null” distributions

- State your null hypothesis in the form “Y does not depend on X”. If this is true, X might as well be random. So:
- Randomise the values of X (keep Y in order if you have other variables not being tested)
  - To randomise, sample *without* replacement (the default in `sample()`)
- Calculate your statistic from the randomised “data” and store it in a suitable array...
- Repeat lots of times (999? 9999? 99999?). E.g.:

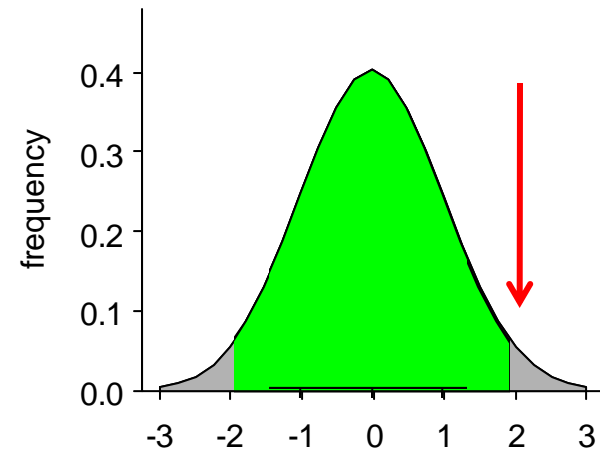
```
> stats.ran <- matrix(NA, nrow=999, ncol=2)
> for(r in 1:999) {
+   X.ran <- sample(data$X)
+   stats.ran[r,] <- lm(Y ~ X.ran)$coef }
```

# DIY “null” distributions

- Now find some quantiles of your “null” distribution to see what values are most likely:

```
> quantile( stats.ran[1,],  
+ probs=c(0.005,0.025,0.975,0.995) )
```

- and compare your true statistic to these.
- One-tailed or two-tailed test?



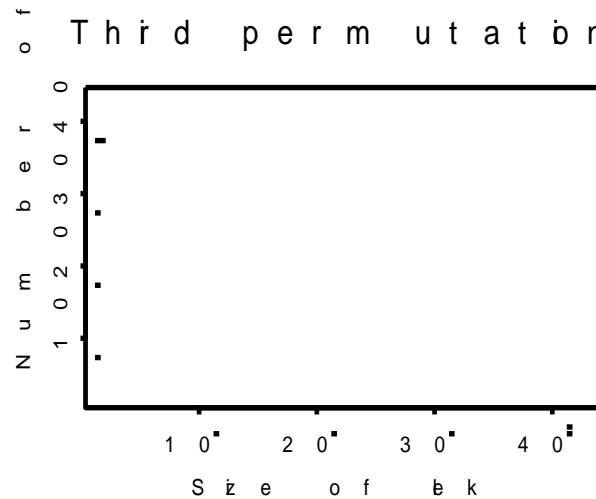
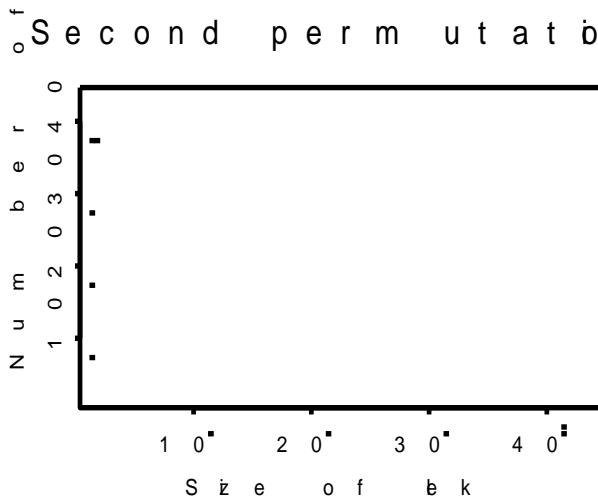
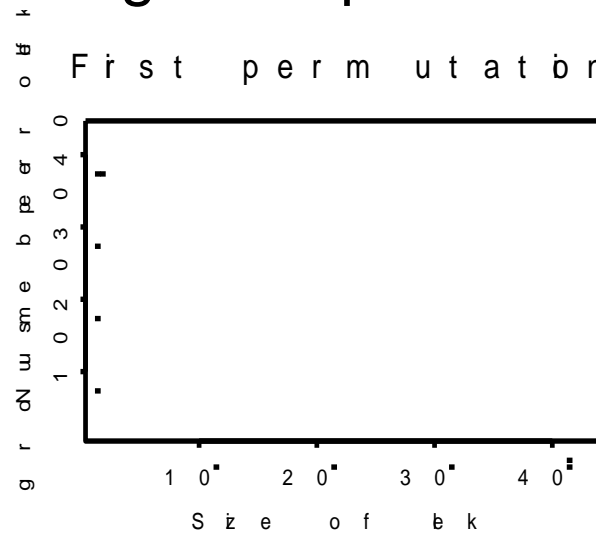
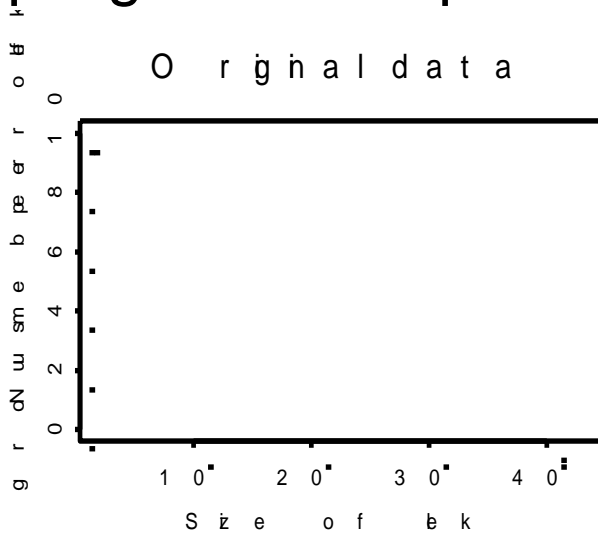
# Example: Permutation test for association

<i>Lek No.</i>	<i>Lek Size</i>	<i>Number of birds</i>	<i>Permutation</i>	<i>Permutation 2</i>	<i>Permutation 3</i>
1	5	0	5	43	24
2	8	1	7	33	19
3	24	3	8	5	34
4	10	1	2	8	6
5	3	1	19	3	12
6	2	0	6	2	2
7	5	0	10	6	33
8	22	8	3	2	4
9	43	10	2	8	7
10	4	4	5	5	2
11	2	1	33	10	43
12	6	5	12	34	8
13	6	0	24	6	3
14	33	8	43	24	10
15	3	0	3	4	22
16	12	1	22	7	6
17	8	1	34	3	5
18	7	0	6	22	8
19	19	1	8	12	3
20	34	5	4	19	5
	<b>Corr=r</b>	<b>0.802</b>	<b>-0.206</b>	<b>0.018</b>	<b>-0.041</b>

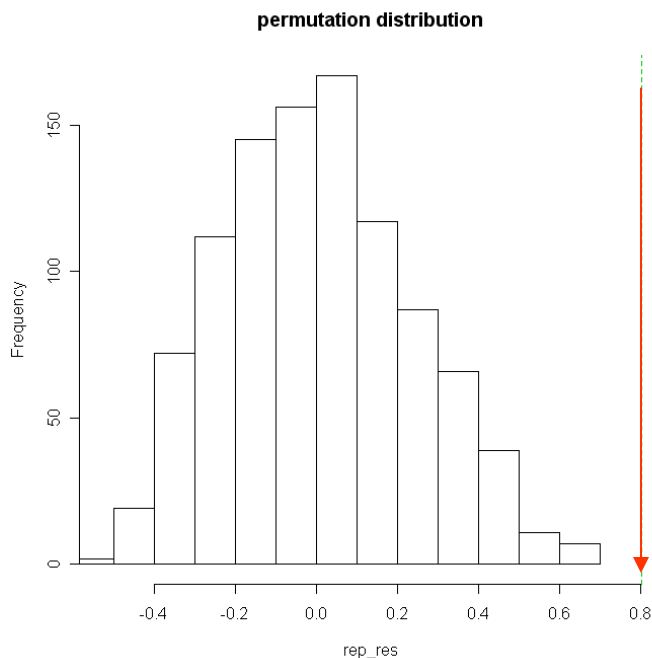


# Example: Permutation test for association

Sampling without replacement gives a permuted sample.



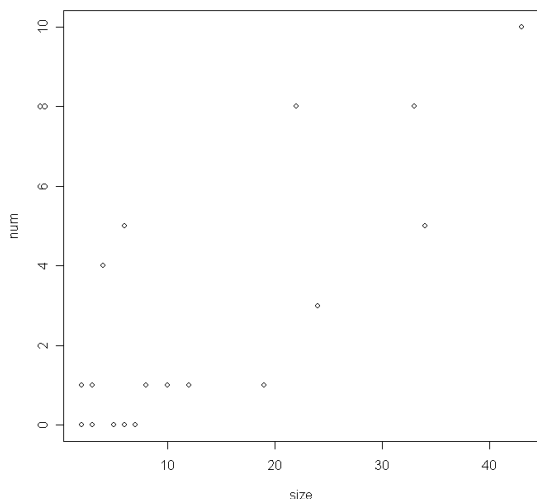
# Example: Permutation test for association



*Observed correlation  $r = 0.802$*

1000 permutations give null distribution of correlation coefficient, all smaller than observed

$P < 0.001$  that observed correlation would arise by chance



Pearson correlation of rank(size) and rank(num) = 0.654  
P-Value = 0.002 (this is Spearman's rank correlation)

Pearson correlation of s and n = 0.802  
Bootstrap  $P$ -Value < 0.001

# Practicals

- Day 5 P6\_resamp
- *This afternoon's P7\_power is hardly a practical, likely to take only a few minutes. Use the rest of the sessions as you want.*