> “The incentive for contemplating a scientific hypothesis is that through it we may achieve an economy of thought in the description of events, enabling us to enunciate laws and relations of more than immediate validity and relevance.”
> Edwards (1992)

## Hypotheses

Up to this point we have really only learned some of the description and implementation issues in statistics.
This will continue, but contemplation of hypotheses is at the core of both statistics and science in general.

## Falsifiability

A central concept in scientific hypothesis testing is that of falsifiability.
In essence this concept states that:

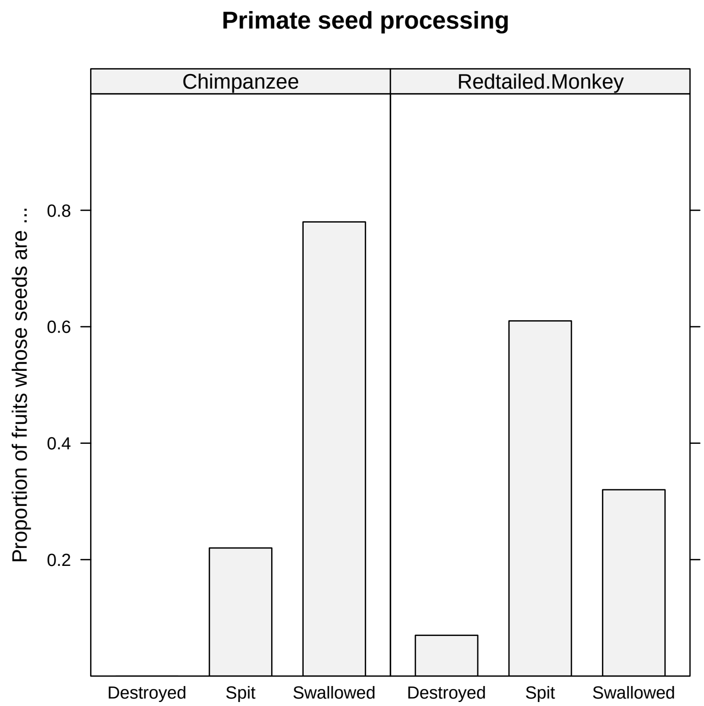
- hypotheses cannot be proven
- science progresses by repeatedly “chipping away” at hypotheses. The resistant ones remain as knowledge.
- hypotheses can be ranked in terms of falsifiability.
- if a proposition is not falsifiable, it is not a scientific hypothesis

Falsifiability was proposed formally by Popper (1959).

## Structure of a hypothesis test

- A model of the real world is developed.
- observed data are compared to that model
- if differences are deemed large enough, the original model may be rejected.

## Seed handling and scarification




Lambert (1999)

### Null Expectation

---

Hypotheses tests begin with null expectations. These represent the 'model of the real world' outlined above.

- Null expectations should be realistic (e.g. seed handling same)
  - avoid trivial expectations (testing the treatment)
  - avoid specifying expectations that apriori you know are not acceptable
- If these null expectations cannot explain observations, alternative expectations must be sought.

### $H_0$

---

These null expectations are generally called the *null hypothesis* , written as $H_0$.

- The null hypothesis is used to make predictions that are compared to the experimental results.
  - Typically, the difference is measured through one or more *test statistics*
- The null hypothesis can be rejected or not. That is the only outcome of an experiment.

- Failing to reject the null does not mean that the null is accepted.

$H_1$

---

- The null is usually the complement of the question in which the experimenter may be interested. The experimenters question may be phrased as an alternative hypothesis, $H_1$.
- Nevertheless, rejecting the null does not imply accepting any of the alternatives.
- In the situation where the null is rejected, it may be worth further examining some $H_1$ and perhaps treating them as $H_0$ in future experiements.

# Part I
# Deviation from expectation under the null

### What are the roles of random effects

---

Random noise is a guarantee when dealing with real-world data
There are two broad sources

- Measurement error
- Natural variation: few biological experiments are immune to:
  - genotypic variation
  - environmental variation

Even if measurement error can be reduced to zero, some random noise will remain in observed data.

### Sampling distributions

---

The expectations of random distributions for many classical tests in statistics are known. They are known as sampling distributions

- Represent infinite repetitions of null process
- Can be used to predict the outcome of these repeated processes

- test statistics are compared to sampling distributions to decide how much random error is tolerable in a particular hypothesis test.

### Deciding on tolerable error

---

This is a *personal decision* with which each scientist has to come to terms.

> "I believe that each scientist and interpreter of experimental results bears ultimate responsibility for his own concepts of evidence and his own interpretation of results." A. Birnbaum (1962)

However, there are accepted norms. **These are guidelines**

### Levels of acceptable error

---

Generally levels of acceptable error are set at:

- 1%
- 5%
- 10%

### Single observation

---

If you are interested in whether a particular number $x$ comes from a normal distribution

- $H_0$: $x\ N(0,1)$
- $H_1$: $x!\ N(0,1)$

### Testing single obs (cont)

---

- Decide on your acceptable level of error
- Calculate the range of values under the null that are possible.
- fail to reject if the sample mean falls in this range

### Simulated example

---

```
> samp <- rnorm(1, 0, 1)
> low.bound <- qnorm(0.025, 0, 1)
> upper.bound <- qnorm(0.975, 0, 1)
> low.bound
```

```
[1] -1.959964

> samp

[1] -0.6394252

> upper.bound

[1] 1.959964
```

**Simulated example (cont.)**

---

```
> samp1 <- rnorm(1, 3, 1)
> low.bound

[1] -1.959964

> samp1

[1] 2.804145

> upper.bound

[1] 1.959964
```

# 1 Types of Error

**Error in Hypothesis testing**

---

There are two classes of error that could occur when testing hypotheses. They are summarized in this table:

| Statistical Decision | Truth | |
|---|---|---|
| | $H_0$ true | $H_0$ false |
| Reject | Type I ($\alpha$) | Correct |
| Fail to Reject | Correct | Type II ($\beta$) |

# 2 Estimation

**Estimation**

*Estmation* is the process of coming up with a reasonable guess at a population parameter, **and characterizing your confidence in the guess**
Parameters that are estimated include mean and variance and other summary statistics that we have seen. Estimates have:

- an estimated value
- an estimate of error

There are many ways to estimate a quantity and its error. We will take the different approaches as they come along.

## 2.1 Estimating means

**Estimating means**

The value of a population estimate of the mean is equal to the arithmetic mean of the sample. This estimator is *unbiased* and has a well known sampling distribution (central limit theorem).
The estimate of confidence in the estimate of a mean is known as the standard error of the mean (SEM).

$$SEM = s_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

Actually the SEM is the standard deviation of the estimate of the mean.

**Confidence intervals**

Confidence intervals provide an interval estimate of a parameter. They say that the true value of the parameter falls between lower ($L_1$) and upper ($L_2$) confidence bounds with XX ($\alpha$) certainty.

$$P\left[-1.96 \leq \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \leq 1.96\right] = 0.95$$

$$P\left[\bar{Y} - \frac{1.96\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + \frac{1.96\sigma}{\sqrt{n}}\right] = 0.95$$

## 2.2 Unkwown parametric $\sigma$

**What if we don't know $\sigma$**

---

If the parametric value of $\sigma$ is unknown, we have to use a different sampling distribution to obtain confidence intervals.
That distribution is the $t$ distribution

**$t$-distribution**

---

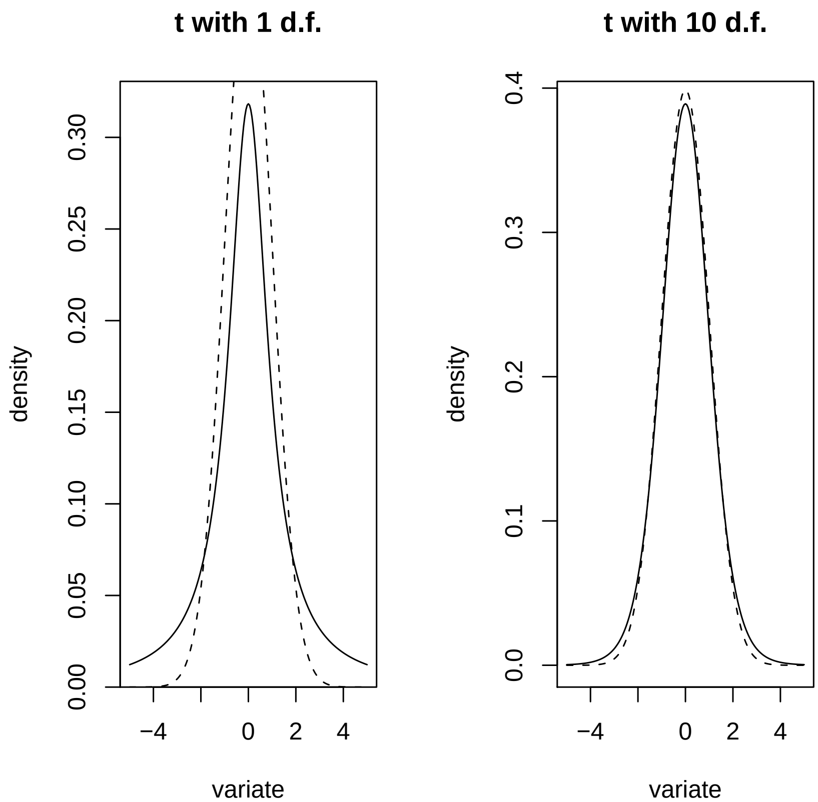
The $t$ distribution is in many ways similar to the normal distribution.

- continuous
- symmetric
- $[-\infty, \infty]$

The final shape of the $t$ distribution is determined by the degrees of freedom. As this number increases, the $t$ distribution approximates the normal. The $t$-distribution is based in part upon a sample variance. The degrees of freedom are $n - 1$

**$t$-distribution**

---




**Confidence intervals based upon $t$ dist**

---

$$L_1 = \bar{Y} - t_{\alpha[df]} s_{\bar{Y}}$$

$$L_2 = \bar{Y} + t_{\alpha[df]} s_{\bar{Y}}$$

Where $df$ is equal to n-1

# References


Edwards, A. W. F. (1992). *Likelihood*, expanded edn, Johns Hopkins University Press, Baltimore, Maryland.

Lambert, J. (1999). Seed handling in chimpanzees (pan troglodytes) and redtail monkeys (cercopithecus ascanius): Implications for understanding hominoid and cercopithecine fruit-processing strategies and seed dispersal, *American Journal of Physical Anthropology* **109**: 365–386.

Popper, K. R. (1959). *The logic of scientific discovery*, Basic Books, New York.