Chapter 5
Hypothesis testing
Goodness-of-fit test

Hypotheses

- Basic components of scientific theories
- Formulated as *universal statements* (concerning all objects of a given class)
  - E.g. **All** mammals are viviparous, **All** plants contain chlorophyll
  - Informative about the members of the given class (even without knowledge of a specific case): A lion is a mammal -> it is also viviparous (
- Must be falsifiable = sensitive to empirical observations, which may demonstrate it is false
- A single contradictory observation can be used to reject a hypothesis no matter how many observations support it
  - Empirical singular statement – not informative beyond the
- No hypothesis can be demonstrated universally true
- Science is a process of rejecting old hypotheses and replacing them with new ones which explain reality better

Platypus is an oviparous mammal.

Rafflesia arnoldii and *Lathraea squamaria* are parasitic plants containing no chlorophyll.

Statistical hypothesis testing

- Empirical evidence mostly comes from measurements of quantitative data
- Statistics quantifies the disagreement between a **null hypothesis (H0)** and empirical observations
  - Measured by a test statistic
    - E.g. χ^2, t, F-ratio
  - Test statistics have a known probability distribution
    - Allows computing the **p-value** = Type I Error probability (probability of making an error if rejecting the H0)
    - Acceptable error typically 0.05 (5%)
    - Shape of the distribution depends also on the **degrees of freedom**
      - These depend on the complexity of the system and/or number of observations

Possible outcomes of hypothesis testing by statistical tests. H_0 = null hypothesis

		Reality	
		H_0 is true	H_0 is false
Our Decision	Reject H_0	Type I Error	Ok
	Not reject H_0	Ok	Type II Error

- Classical biostatistics quantifies precisely only the **Type I Error**
- Hypotheses cannot be confirmed!
- A non-significant test means that empirical evidence is not sufficient to reject the H0

Pattern detection

- Special cases when H0 = 0 (no effect, independence)
- Generic null hypotheses can be converted to pattern detection problems by subtracting hypothesis expectations from observed data
- Generally easy to compute
- Biology is still largely exploratory science
  - many biological problems have the pattern-detection nature

Goodness of fit test

- Suitable for frequencies (count data)
- Formula: $\chi^2 = \sum \frac{(O-E)^2}{E}$
  - O = observed frequencies = empirical counts
  - E = expected frequencies = counts predicted by the null hypothesis
- Degrees of freedom = number of categories
- Report: χ^2_{df} = value, p = value
- Colloquially called χ^2- test (chi-square square)