t- and F-tests

Testing hypotheses

Overview

- Distribution& Probability
- Standardised normal distribution
- t-test
- F-Test (ANOVA)

Starting Point

- Central aim of statistical tests:
  – Determining the likelihood of a value in a sample, given that the Null hypothesis is true: $P(value|H_0)$
    - H_0: no statistically significant difference between sample & population (or between samples)
    - H_1: statistically significant difference between sample & population (or between samples)

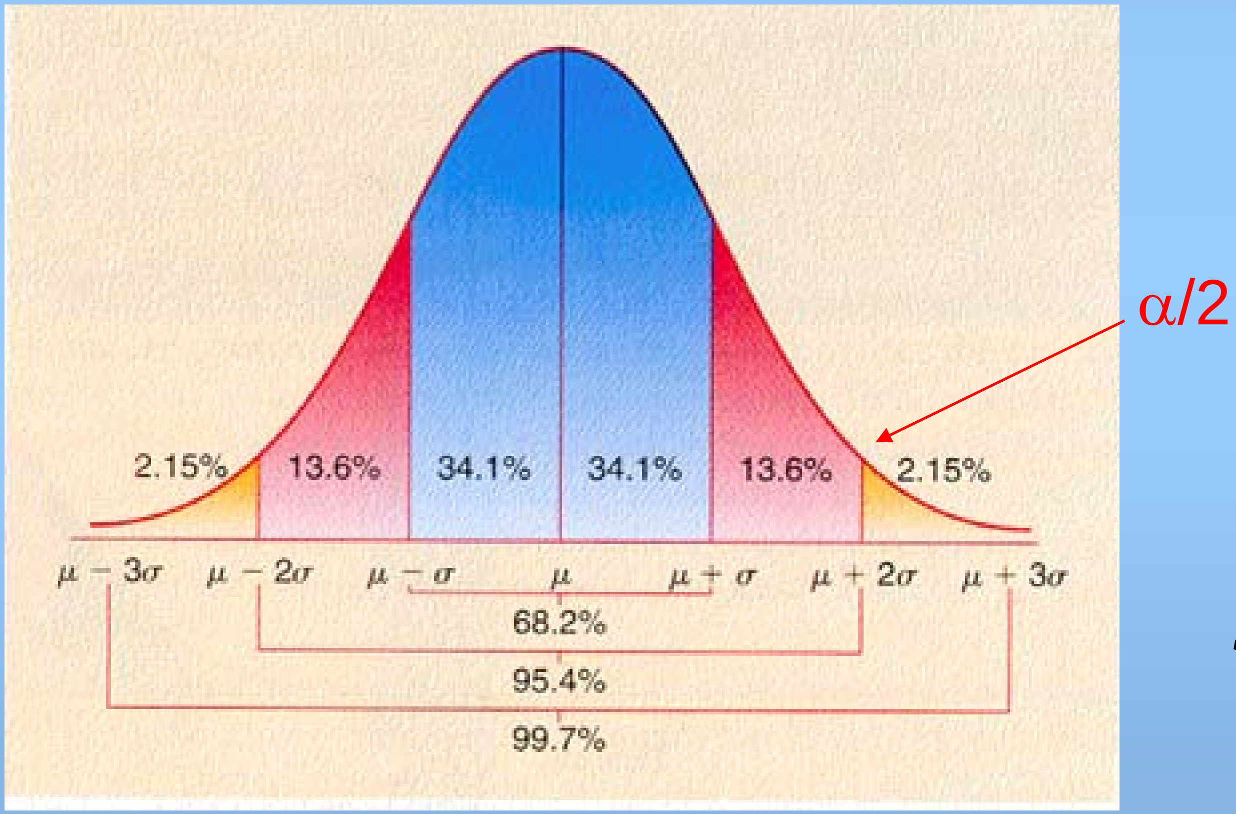
  – Significance level: $P(value|H_0) < 0.05$

Types of Error

		Population	
		H_0	H_1
Sample	H_0	$1-\alpha$	β-error (Type II error)
	H_1	α-error (Type I error)	$1-\beta$

Distribution & Probability

If we know s.th. about the distribution of events, we know s.th. about the probability of these events



$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

Standardised normal distribution

Population

$$z = \frac{\bar{x} - \mu}{\sigma}$$

Sample

$$z_i = \frac{x_i - \bar{x}}{s}$$

$$\bar{x}_z = 0$$

$$s_z = 1$$

- the z-score represents a value on the x-axis for which we know the p-value

- 2-tailed: z = 1.96 is 2SD around mean = 95% → ‚significant‘
- 1-tailed: z = +-1.65 is 95% from ‚plus or minus infinity‘

t-tests:

Testing Hypotheses About Means

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$$

$$t = \frac{differences\_between\_sample\_means}{estimated\_standard\_error\_of\_differences\_between\_means}$$

Degrees of freedom (df)

- Number of scores in a sample that are free to vary

- n=4 scores; mean=10 → df=n-1=4-1=3
  - Mean= 40/4=10
  - E.g.: score1 = 10, score2 = 15, score3 = 5 → score4 = 10

Kinds of t-tests

Formula is slightly different for each:

- Single-sample:
  - tests whether a sample mean is significantly different from a pre-existing value (e.g. norms)
- Paired-samples:
  - tests the relationship between 2 linked samples, e.g. means obtained in 2 conditions by a single group of participants
- Independent-samples:
  - tests the relationship between 2 independent populations
  - formula see previous slide

Independent sample t-test

Number of words recalled

Group 1	Group 2 (Imagery)
21	22
19	25
18	27
18	24
23	26
17	24
19	28
16	26
21	30
18	28
mean = 19	mean = 26
std = sqrt(40)	std = sqrt(50)

df = $(n_1-1) + (n_2-1) = 18$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{19 - 26}{1} = -7$$

$$t_{(0.05,18)} = \pm 2.101$$

$$t > t_{(0.05,18)}$$

→ Reject H_0

Bonferroni correction

- To control for false positives:

$$p_c = \frac{p}{n}$$

- E.g. four comparisons:

$$p_c = \frac{0.05}{4} = 0.0125$$

F-tests / Analysis of Variance (ANOVA)

T-tests - inferences about 2 sample means

But what if you have more than 2 conditions?

e.g. placebo, drug 20mg, drug 40mg, drug 60mg

Placebo vs. 20mg	20mg vs. 40mg
Placebo vs 40mg	20mg vs. 60mg
Placebo vs 60mg	40mg vs. 60mg

Chance of making a type 1 error increases as you do more t-tests

ANOVA controls this error by testing all means at once - it can compare *k* number of means. Drawback = loss of specificity

F-tests / Analysis of Variance (ANOVA)

Different types of ANOVA depending upon experimental design (independent, repeated, multi-factorial)

Assumptions

- observations within each sample were independent
- samples must be normally distributed
- samples must have equal variances

F-tests / Analysis of Variance (ANOVA)

$$t = \frac{\text{obtained difference between sample means}}{\text{difference expected by chance (error)}}$$

$$F = \frac{\text{variance (differences) between sample means}}{\text{variance (differences) expected by chance (error)}}$$

Difference between sample means is easy for 2 samples:

(e.g. $\bar{X}_1$=20, $\bar{X}_2$=30, difference =10)

but if $\bar{X}_3$=35 the concept of differences between sample means gets tricky

F-tests / Analysis of Variance (ANOVA)

Solution is to use variance - related to SD

$$Standard\ deviation = \sqrt{Variance}$$

E.g.

Set 1	*Set 2*
20	*28*
30	*30*
35	*31*
$s^2=58.3$	$s^2=2.33$

These 2 variances provide a relatively accurate representation of the size of the differences

F-tests / Analysis of Variance (ANOVA)

Simple ANOVA example

Total variability

Between treatments variance

Measures differences due to:

1. Treatment effects
2. Chance

Within treatments variance

Measures differences due to:

1. Chance

F-tests / Analysis of Variance (ANOVA)

$$F = \frac{MS_{between}}{MS_{within}}$$

When treatment has no effect, differences between groups/treatments are entirely due to chance. Numerator and denominator will be similar. *F*-ratio should have value around 1.00

When the treatment does have an effect then the between-treatment differences (numerator) should be larger than chance (denominator). *F*-ratio should be noticeably larger than 1.00

F-tests / Analysis of Variance (ANOVA)

Simple independent samples ANOVA example

	Placebo	Drug A	Drug B	Drug C
Mean	1.0	1.0	4.0	6.0
SD	1.73	1.0	1.0	1.73
n	3	3	3	3

$F(3, 8) = 9.00$, $p<0.05$

There is a difference somewhere - have to use post-hoc tests (essentially t-tests corrected for multiple comparisons) to examine further

F-tests / Analysis of Variance (ANOVA)

Gets more complicated than that though…

Bit of notation first:

An independent variable is called a *factor*

e.g. if we compare doses of a drug, then dose is our factor

Different values of our independent variable are our *levels*

e.g. 20mg, 40mg, 60mg are the 3 levels of our factor

F-tests / Analysis of Variance (ANOVA)

Can test more complicated hypotheses - example 2 factor ANOVA (data modelled on Schachter, 1968)

Factors:

1. Weight - normal vs obese participants
2. Full stomach vs empty stomach

Participants have to rate 5 types of crackers, dependent variable is how many they eat

This expt is a 2x2 factorial design - 2 factors x 2 levels

F-tests / Analysis of Variance (ANOVA)

Mean number of crackers eaten

	Empty	Full	
Normal	22	15	= 37
Obese	17	18	= 35
	= 39	= 33	

Result:

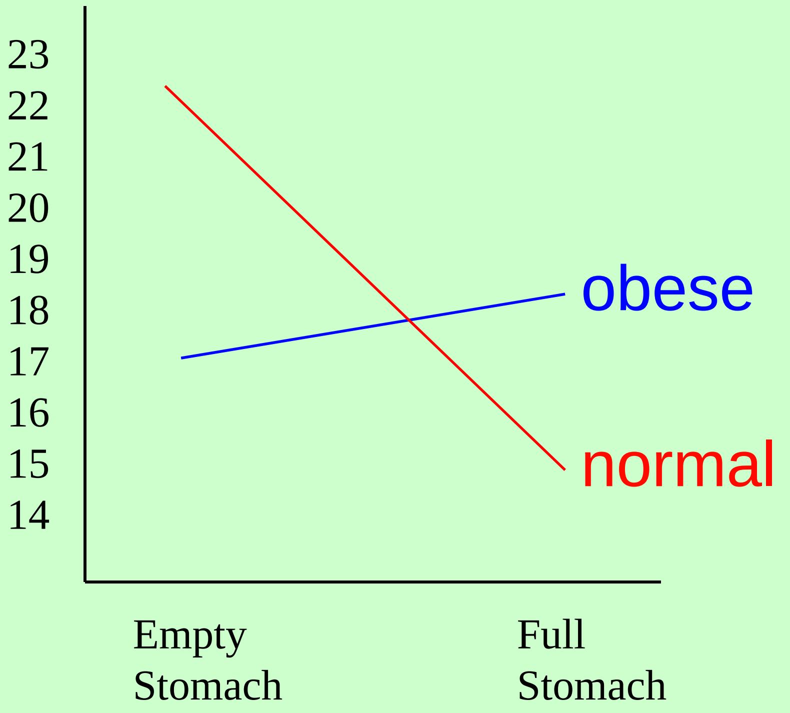
No main effect for factor A (normal/obese)

No main effect for factor B (empty/full)

F-tests / Analysis of Variance (ANOVA)

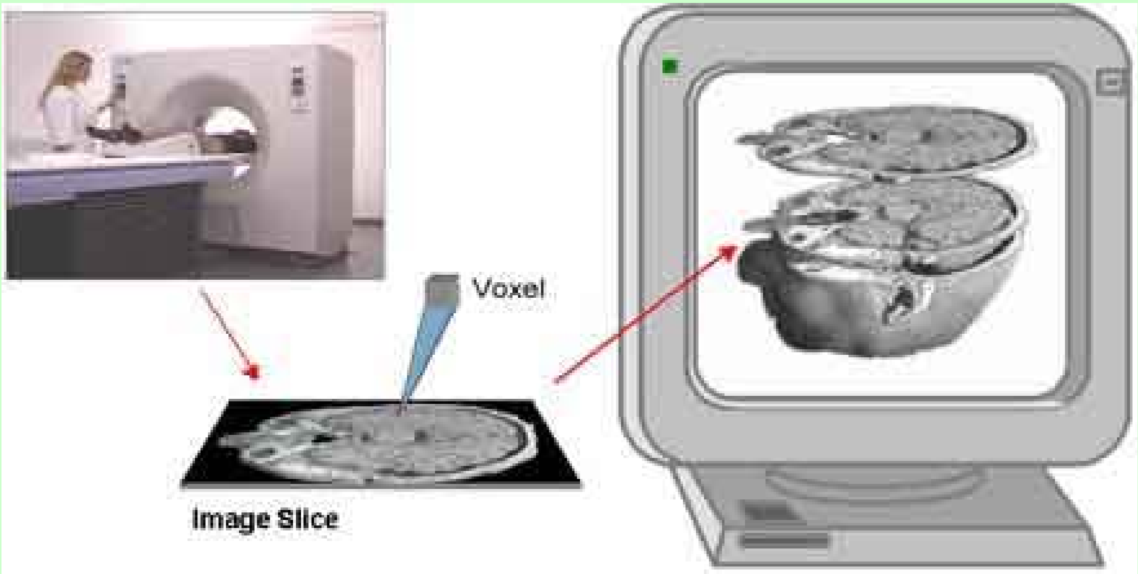
Mean number of crackers eaten

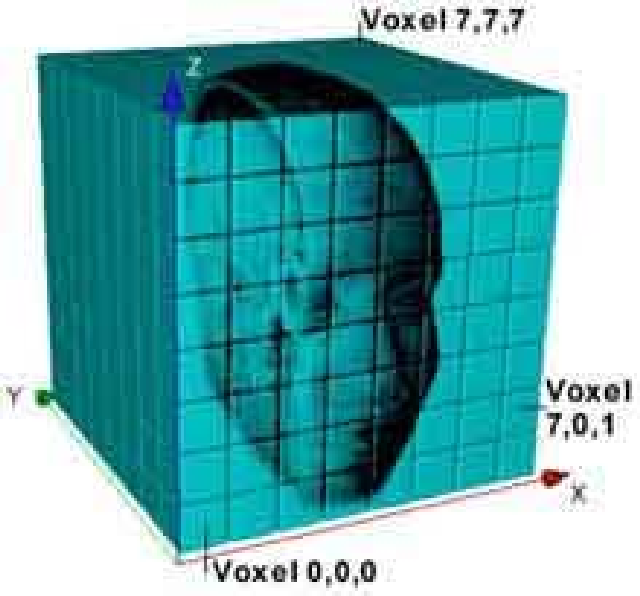
	Empty	Full
Normal	22	15
Obese	17	18



F-tests / Analysis of Variance (ANOVA)

Application to imaging…

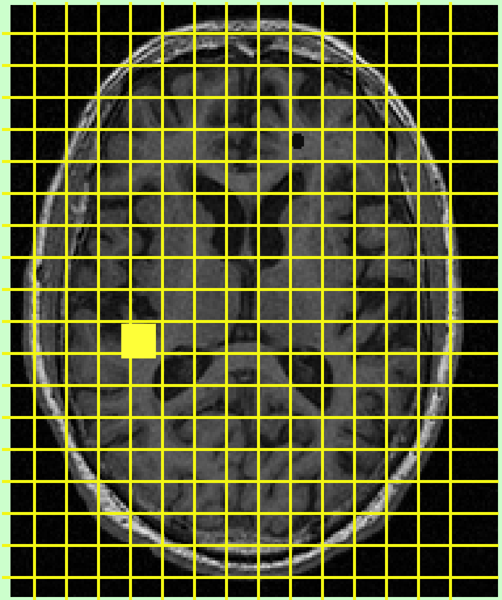


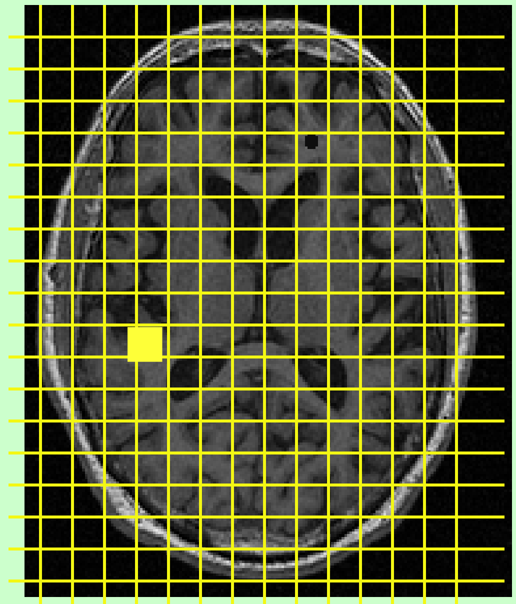


F-tests / Analysis of Variance (ANOVA)

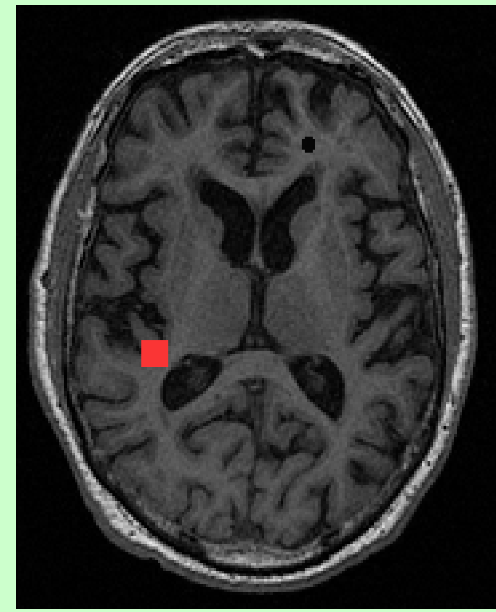
Application to imaging…

Early days => subtraction methodology => T-tests corrected for multiple comparisons





=



e.g. Pain
Visual task

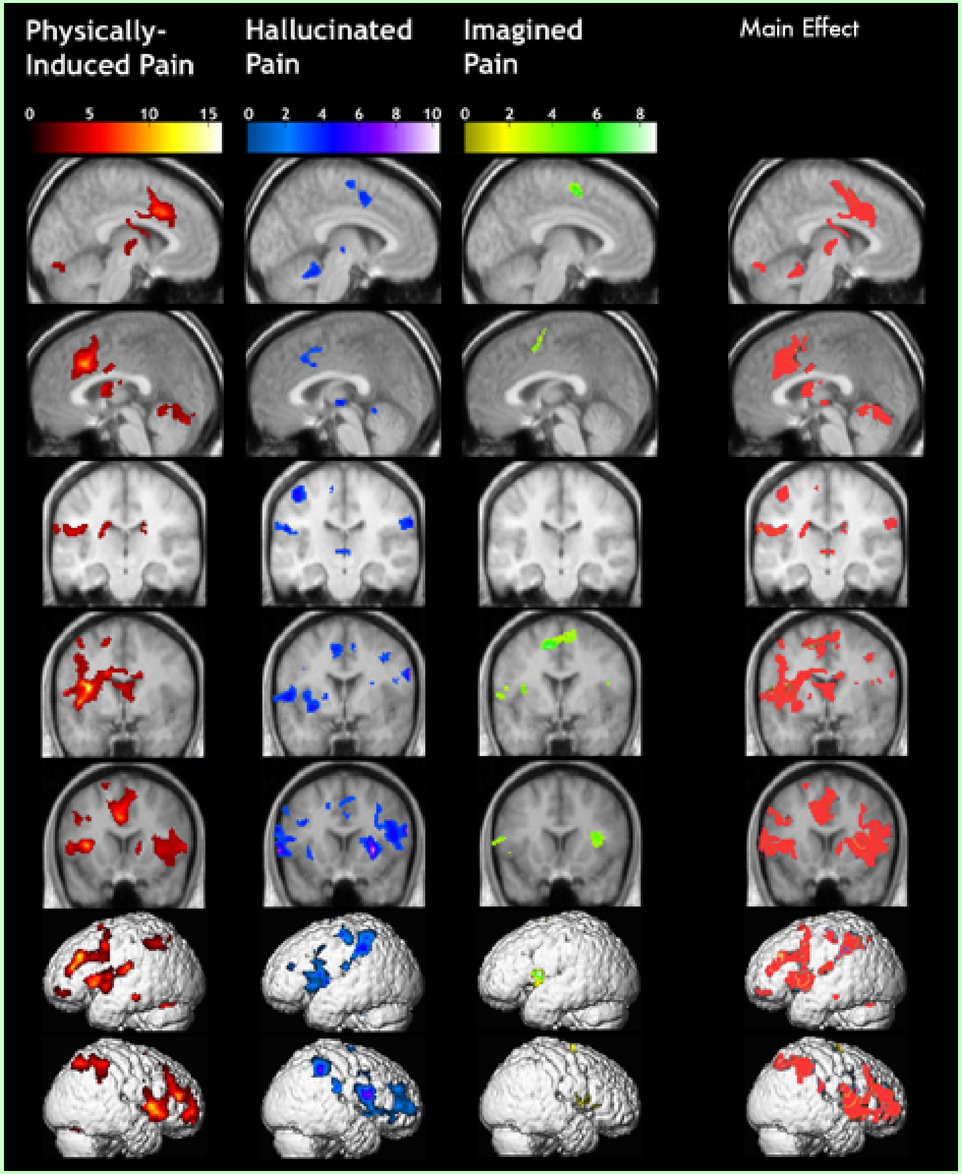
\-

Appropriate rest condition

=

Statistical parametric map

F-tests / Analysis of Variance (ANOVA)



This is still a fairly simple analysis. It shows the main effect of pain (collapsing across the pain source) and the individual conditions.

More complex analyses can look at interactions between factors

Derbyshire, Whalley, Stenger, Oakley, 2004

References

Gravetter & Wallnau - Statistics for the behavioural sciences

Last years presentation, thank you to:

Louise Whiteley & Elisabeth Rounis

http://www.fil.ion.ucl.ac.uk/spm/doc/mfd-2004.html

Google