- BETA should be lower than ALPHA if the cost of the trial is higher than its benefit to knowledge (i.e. looking at a drug and its side effects).

One/single sample *t* test

Collecting one set of data and comparing it to another value (i.e. collecting heights and comparing it to an average height)

Assumes that the sample is random and following a normal distribution

We can test whether the population mean differs from a particular value ($\mathbf{H}_0$: $\mu = \mu_0$) by calculating the following statistic from the sample: $t = (m - \mu_0)/SE$ =mean – reference value/SE

A sample statistic that is used to assess a null hypothesis is called a test statistic.

This particular one is called the one-sample *t*-test statistic.

(Theory shows that it can be compared against values from a *t*-distribution with n-1 degrees of freedom (d.f.) to make a test decision - hence the name.)

Reporting a one sample t test: t (490) = -7.241; $p < .001$

Testing a one sample proportion

When you look at the probability of a category against a pre-specified value

Or it might be when you want to see if a group is under-representative.

Conduct a one-sample Chi-squared test (non-parametric). This is an alternative test allowing tests of proportions rather than means.

Analyse → non-parametric → legacy dialogues → chi-squared

Non-parametric as proportions are not normally distributed.

Assumption is that each category will have more than a value of 5 (i.e. 3 males and 200 females)

If the sample is large enough you can use a binomial test which is an exact test

Asymptotic or a Monte-Carlo test when the sample is small.

Preview from Notesale.co.uk
Page 11 of 29

Measures of effect

When comparing groups, you look at exposure – outcome

Measures of effect = indices that quantify the strength, and direction, of the E – O association.

Measurement scale	Effect/impact index
Interval (continuous outcomes)	- Mean Difference - Median Difference
Count (incidence rates)	- Incidence Rate Ratio (IRR)
Binary (prevalence or incidence proportions)	- Risk Difference (RD) - Relative Risk (RR) - Odds Ratio (OR)

Independent samples *t*-test

Inferences assume both groups are samples from normal distributions.

The difference in population means can be estimated by the difference between sample means: $d=m_1-m_2$

SE(d) = standard error of difference between means

- Can be estimated from the within group standard deviations

The independent samples *t*-test (equal variances assumed) compares a *t-statistic (t):* $\frac{d}{SE(d)}$ with a *t*-distribution with *n*-2 degrees of freedom.

95% Confidence interval for a difference between means: $[d - 2 \times SE(d), d + 2 \times SE(d)]$

Test the null hypothesis to investigate whether the means are the same: $\mu_1 = \mu_2$. Possible alternative Hypotheses H_A = Means are different: $\mu_1 \neq \mu_2$ (two-sided) or one population has a larger mean: $\mu_1 > \mu_2$ ($_{one}$-sided), or $\mu_1 < \mu_2$ ($_{one}$-sided)

Two test versions available: one assumes equal variances (homogenous), the other does not.

Levenes test

Levenes test is an inferential statistic used to assess the assumption of equality of variance (homogeneity of variance or homoscedasticity) for a variable calculated for two or more groups.

Levenes test assesses the assumption of equal variance of the two groups

It tests the null hypothesis that the population variances are equal

If the Sig. Value is greater than .05…

- Read from first row. A value greater than .05 means that the variability in your two conditions is about the same. That the scores in one condition do not vary too much more than the scores in your second condition. Put scientifically, it means that the variability in the two conditions is not significantly different.
- Two variances are not statistically different so the variances are equal.
- We do not reject the null hypothesis of equal variances and it is concluded that the assumption of Homoscedasticity is valid

If the Sig. Value is less than or equal to .05…

- Read from second row. A value less than .05 means that the variability in your two conditions is <u>not</u> the same. That the scores in one condition vary much more than the scores in your second condition. Put scientifically, it means that the variability in the two conditions is significantly different. This is a bad thing, but SPSS takes this into account by giving you slightly different results in the second row.
- We reject the null hypothesis of equal variances and it is concluded that there is a difference between the variances in the population.

Preview from Notesale.co.uk
Page 12 of 29

		Suffer from malaise at age 22 years?		
		no	yes	Total
Gender	male	465	26	491
	female	437	72	509
	Total	902	98	

Two way tables allow the calculation of risk and odds.

You can calculate row percentages or column percentages depending on the chi square you want to perform.

You can display them as a percentage of the total population.

****** pay attention in the exam to what the percentages are actually representing ******

Relative risk: $RR = \frac{Risk\ of\ Males\ suffering\ from\ Malaise}{Risk\ of\ Females\ suffering\ from\ Malaise}$

$$= \frac{26}{491} \div \frac{72}{509} = \frac{26 \times 509}{491 \times 72} = 0.374$$

So, there is a 37% less risk of males suffering from Malaise

Odds ratio: OR $= \frac{The\ chance\ of\ Males\ suffering\ Malaise\ compared\ to\ not\ suffering\ Malaise}{The\ chance\ of\ Females\ suffering\ from\ Malaise\ compared\ to\ not\ suffering\ Malaise}$

$$= \frac{26}{465} \div \frac{72}{437} = \frac{26 \times 437}{465 \times 72} = 0.339$$

Interpreting odds:

- OR = 1 = Exposure does not affect the odds of outcome
- OR < 1 = Indicates the exposure is associated with the reduced risk of developing the outcome
- OR >1 = Indicates that the exposure is associated with an increase risk of developing the disease

E.g. if the odds ratio = 1.5 then the odds of males suffering malaise is 1.5 times that of females suffering from Malaise.

Preview from Notesale.co.uk
Page 14 of 29

An odds ratio is a special type of ratio, one in which the numerator and denominator sum to one.

For contingency tables the association between row and the column variable is typically of interest. We may wish to check whether there is any association between the row and column variable.

We can formally test the null hypothesis that the row and column variables are statistically independent.

Statistical independence means that knowing the value of one variable does not provide any information regarding the distribution of the other, or formally for any two categories A and B:

$$Prob(A\ and\ B) = Prob(A) \times Prob(B)$$

For statistical test of independence, a Chi Squared test is used.

Chi squared test

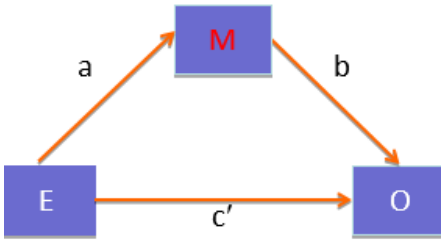
The chi-squared test of independence allows the formal comparison of two categorical (nominal) variables in a sample to test if there is any association between them – either independent groups or when comparing one sample with a reference figure.

The idea behind this test is to compare the observed frequencies with the frequencies that would be expected if the null hypothesis of no association / statistical independence were true.

Mediation

Mediation is a hypothesised causal mechanism by which one variable affects another.

A mediator of the causal effect of exposure on outcome is a variable on the causal pathway from E to O.

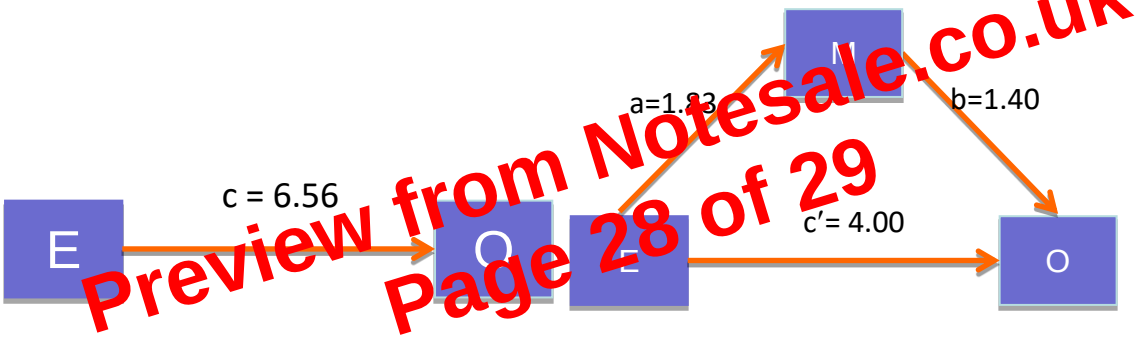


In a mediated model the total causal effect (c) can be split into an indirect and direct part.

In non-mediated models the total effect of exposure on outcome is denoted by path c'. In a mediated model, the total causal effect can be split into an indirect (/mediated) part and a direct (/non-mediated) part.

For a linear regression with a continuous mediator, indirect effect can be obtained by calculating the product of path a and path b (multiplying the pathways)

In regression type models, total effect is the sum of the direct and indirect effects: C = c' + a*b



Preview from Notesale.co.uk
Page 28 of 29

Baron and Kenny Steps

Baron and Kenny (1986) discussed four steps to establish mediation:

1. Test path C (E → O)
   Establish that E is associated with O
2. Test path a (E → M)
   Show that the causal variable is associated with the mediator
3. Test path b (M, controlling for E → O)
   Show that the mediator, adjusted for the causal variable affects the outcome
4. Test path c' (E, controlling for M → O)
   For complete mediation this should be not significant
   For partial mediation c' is smaller than c in absolute value

If indirect effect is the same as the total effect, we have complete mediation. i.e. in the example if the direct effect as 0 it would be complete mediation.