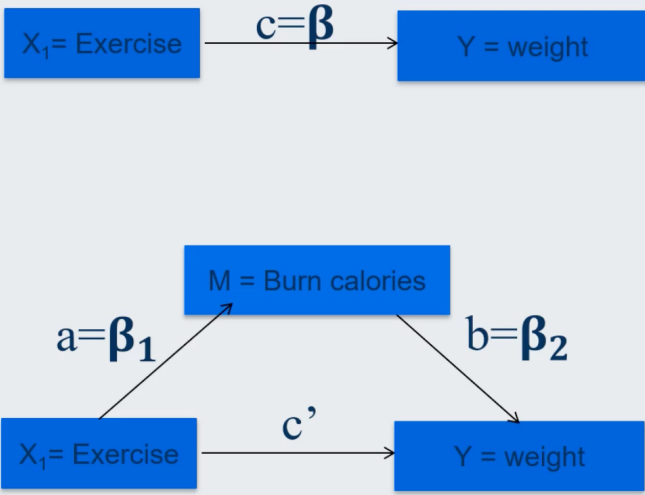


1 sample t A->compare mean->one sample t-test Enter value	-Skewed continuous -Ordinal, interval or discrete	Wilcoxon Signed Rank A->non-p->one sample [fields] var [settings] test type, value
2 independent sample t Levene's test for equal variances P<0.05, not assumed, line 2	Split file: data-split file- organise by groups N* reunite	Mann Whitney U A->non-p->independent sample [fields] var [settings] customise M-W U
2 paired samples t T->compute var Post-pre		Wilcoxon matched pair Signed Rank A->non-p->related sample [fields] var [settings] test type
1 sample χ^2 test A->non-parametric->legacy dialogs-> χ^2 Enter %	-The number of cells with expected frequencies less than 5, are less than 20% -The minimum expected frequency is at the very least 1	1 sample Binomial exact A->non-p->legacy dialogs->binomial?
Pearson χ^2 test - independent A->descriptive->crosstabs Row: gender/exposure Column: outcome [statistics] χ^2 [cells] counts: observed Percentages: row? Odds & Risks [statistics] Risk		Fisher's exact A->descriptive->crosstabs [exact] exact? [statistics] ?
McNemar χ^2 test – paired A->descriptive->crosstabs Row: pre Column: post [statistics] McNemar [cells] counts: observed, total? *pre-yes-total vs post-yes-total More than 2 categories -- McNemar Browker		Binomial exact
Pearson's Correlation A->correlate->bivariate Pearson	-Skewed continuous -ordinal	Spearman's correlation A->correlate->bivariate Spearman
Simple linear regression A->regression->linear [statistics] coefficient intervals (estimates, model fit) Multiple linear regression A->regression->linear [statistics] coefficient intervals (estimates, model fit)	DV - continuous DV - Binary	Binary logistic regression A->regression->binary logistic [options] CI for exp(B) If categorical IV [categorical] (even it has 7 cats) *exp(B) is odds ratio $\ln\left(\frac{\pi}{1-\pi}\right) = L \quad \pi = \frac{1}{1+e^{-L}}$

<p>Assumptions A->regression->linear [plots] tick 3 (histogram, normal p plot, produce all partial plots) ZRESID→Y, ZPRED→X</p> <p>Predict Add a new x value in the table A->regression->linear [save] 'predicted values' unstandardised; 'prediction intervals' mean or individual</p> <p>Categorical IV more than 2 cats T->recode into dif [old and new values]</p> <p>Outliers DFBETA, DFFIT A->regression->linear [save] standardised DfBeta, standardised DfFit New column in table values >1, influential observations</p>		<p>Model is Omnibus tests Explain Nagelkerke R2</p> <p>Model fit: 1 classification table</p> <p>2 Hosmer-Lemeshow goodness of fit [options] H-L goodness of fit p>0.05 good fit</p>
--	--	--

Confound C->X, Y	C as X2
<p>Mediate</p>  <p>4: Estimate of path c': <u>2 different ways:</u></p> <p>i. $c = c' + a * b$ $\beta = c' + \beta_1 * \beta_2$</p> <p>ii. From step 3 model: $Y = \beta_0 + \beta_2 M + \beta_3 X_1 + \epsilon$ $c' = \beta_3$</p>	<p>Total effect c Mediated indirect ab Non-m direct c'</p> <p>Mediation needs a*, b* Complete or partial c'</p> <p>Testing ab: Sobel test test statistics >1.96 reject ab=0</p> <p>Non-p Sobel/ process Check if 0 in CI</p>
<p>Interact/ Modify Create interaction terms: T->compute var 'Target Var' height_x_sex 'numeric expression' height * sex</p>	<p>A->regression->linear</p> <p>Dummy Var T->recode into dif [old and new values]</p>

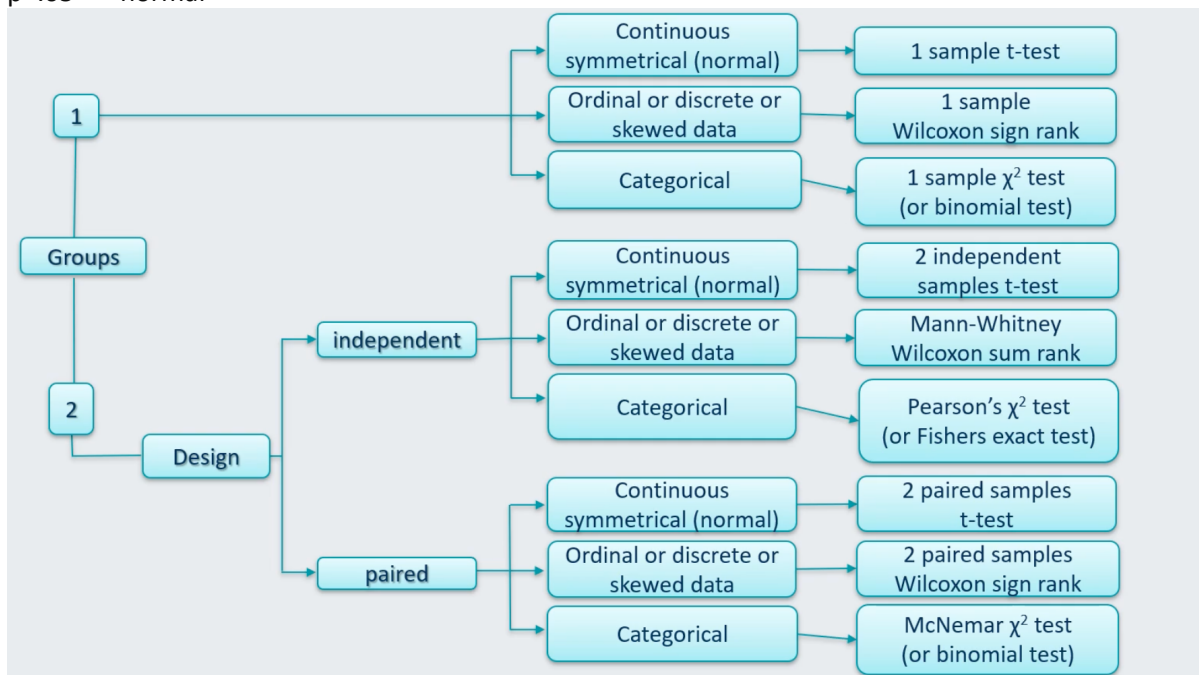
Scatter plots
G->scatter/dot
Simple scatter

Transform-compute var

Check normality
One-sample Kolmogorov-Smirnov test:
A --> nonparametric test --> legacy dialogue --> 1-sample K-S

put the variables interested --> Asymp. Sig. (2-tailed)

>.05 --> normal



0.8-1 very strong, 0.6-0.79strong, 0.4-0.59moderate, 0.2-0.39weak, 0-0.19very weak-none

Percentage v valid percentage x, 'of those who responded' --> valid percent

Pay attention to if there's a % or not

*female or male

卡方 百分比 value 先输 0 对应的, 再输 1

V or adj depends or not independent

If there's 'different' / 'differ' ----- was not significantly different from / did not differ significantly

A weak positive linear relationship

Odds - larger

Regression line: A deterministic equation (no ϵ).

Regression model: A stochastic equation (includes ϵ).

Partial plots - x1&y, x2&y scatter plots

Discrete var: doctor: doctor availability per 100,000 residents

Continuous var: crime – violent crime rate (per 100,000 people)

$(\text{Risk when male}) \div (\text{Risk when female}) = 0.053 \div 0.141 = 0.376$

Those who were male had 0.38 times the risk of having malaise at 22 compared to those who were female.

We can also present risk as a percentage using the following formula % decrease = $(RR - 1) \times 100$

Males had a 62.4% reduction in risk of malaise at 22 compared to females.

the variable 'income' has more than 5 ordered points which are equidistant, thus it is an interval variable with enough points to be treated as continuous (skewed).

有 binary 不选 nominal

Skewed is not bell-shaped

The error terms have [the same variance] irrespective of the values of x. We can inspect this by plotting a [scatterplot] of the [standardised predicted versus standardised residual] values. The plot showed [no pattern].

小数点后 0 不要省

x times more likely / likelihood ---> odds ratio / $\exp(B)$

Model shows that an increase in age by one year is associated with an increased likelihood of adherence.

Binary logistic 默认用 0/reference, 注意提问方式, odds ratio, 换一个为倒数, 换两个不变

量表变量类型不能选 continuous, 选 ordinal / interval? ?

X is a interval variable but because the variable has more than 5 points on a response scale, we decide to treat it as a continuous variable for our analysis. ??

Without figure—ordinal, with figure---might be interval, with figure and over 7 cats---might be treated as continuous