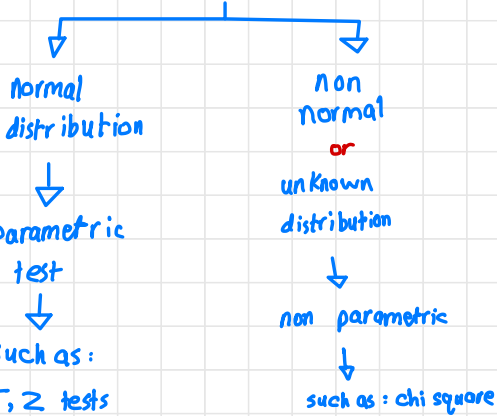


population distribution



Assumptions: two levels
 1) independent "dichotomous" or dependent "continuous"
 2) same variances "homogeneity"
 3) random sampling.

1) no assumptions about data distribution, "distribution-free"
 2) can be with small sample.

3) ONLY: nominal and ordinal data.

4) good with outliers.

5) no overlap

4) no Overlap!!!
 cf dependent + testable overlap.
 -more Powerful

Test of Normality:

we assume that our data is non-normally distributed H_1 .
 $\Rightarrow \alpha > P\text{-value} \Rightarrow H_0$ is rejected, H_1 is accepted.

* Chi square:

- nominal data, sample > 20 .
- establish the level of significance " α "
- formulate the "statistical" hypothesis = H_0 and null hypothesis H_1
To alternative = statistical hypothesis
- calculate the statistical: " χ^2_{calc} "
- determine the degree of freedom.
- compare the static against critical " χ^2_{cv} "
- not Prove causality.
- Actual count not percentage data.

The chi square test can only be used on data that has the following characteristics:

The data must be in the form of frequencies

The frequency data must have a precise numerical value and must be organised into categories or groups.

The expected frequency in any one cell of the table must be greater than 5.

The total number of observations must be greater than 20.

Degree	Frequency	Relative Frequency/Percentage
High School	2	0.050 5.0
Bachelor's	7	0.175 17.5
MBA	20	0.500 50.0
Master's	3	0.075 7.5
Law	4	0.100 10.0
PhD	4	0.100 10.0
	40	

Level of Measurement	Sample Characteristics					Correlation
	1 Sample	2 Sample		K Sample (i.e., >2)		
		Independent	Dependent	Independent	Dependent	
Categorical or Nominal	χ^2	χ^2	Macnamar's χ^2	χ^2	Cochran's Q	
Rank or Ordinal		Mann Whitney U	Wilcoxin Matched Pairs Signed Ranks	Kruskal Wallis H	Friendman's ANOVA	Spearman's rho
Parametric (Interval & Ratio)	z test or t test	t test between groups	t test within groups	1 way ANOVA between groups	1 way ANOVA (within or repeated measure)	Pearson's r
		Factorial (2 way) ANOVA				

$H_1 \Rightarrow$ dependent \Rightarrow المقاطعة
 إذا كانت العلاقة بين الأختلافات، ففرضية H_0 هي أننا افترضنا أنهم مرتبطين ببعضى.

إذا H_0 ، فرب تكون عكس H_1 ، يعني رب يكون باله H_0 إنه ففة الاختلاف غير مرتبطة بالأختلافات، ففرضية H_0 هي المقاطعة.

إذا كانت $\alpha < P\text{-value}$ \Rightarrow معناه إنه H_1 خاطئ \Leftarrow فيقبل H_0
 إذا كانت $\alpha > P\text{-value}$ \Rightarrow معناه إنه H_1 صحيحة \Leftarrow ونفى H_0

* data لا تتصلب
 * degree of freedom
 * $P\text{-value} < \chi^2 < F_{\alpha} < F_{\alpha} < r < H_0 < \chi^2 < N^*$
 * $P\text{-value} < \chi^2 < F_{\alpha} < F_{\alpha} < r < H_0 < \chi^2 < N^*$

$F_e \Rightarrow$ expected frequency = $\frac{\text{row} \times \text{column}}{\text{total}} = \frac{(a+b)(a+c)}{a+b+c+d}$

a	b	a+b
c	d	a+b+c+d
a+c	a+b+c+d	

$\chi^2 \Rightarrow$ chi square = $\sum \frac{(F_o - F_e)^2}{F_e}$

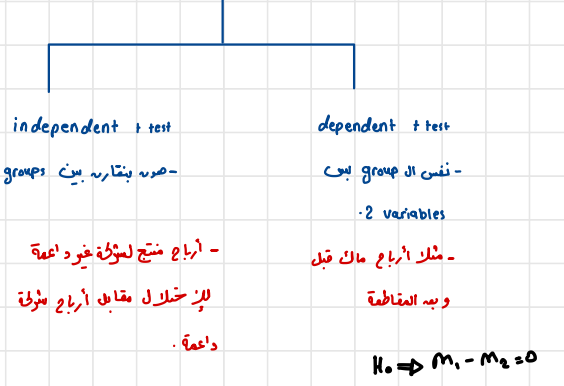
$r = df = (R-1)(C-1)$

* $\chi^2_{\text{calc}} > \chi^2 \Rightarrow$ reject $H_0 \Rightarrow$ **يتم على مثالنا السابق انه الي بوطن**
 من مال له المقابلة هو نفس ما عنده اختلاف.

T test:
 "William Gosset test"

- T test is used for comparing mean value of two groups.

* بال H_0 يفرض انه ال data لا 2 groups ايضا كانت نفس ال distribution فبا
 الهم نفس ال means \Rightarrow كل ما زاد ال test + كل ما زاد الفرق بين ال means كل
 ما بنقص ال H_0 الكثر



1) independent t test:
 $H_0 \Rightarrow M_1 - M_2 = 0$
 $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}}$
 $df = n_1 + n_2 - 2$
 * levene's test of equals of variance
 * non sig \rightarrow reject H_0

2) dependent t test:
 $t_d = \frac{\bar{D}}{\sqrt{\frac{SD_d^2}{n}}}$
 $df = n - 1$
 $\alpha > P$ reject H_0