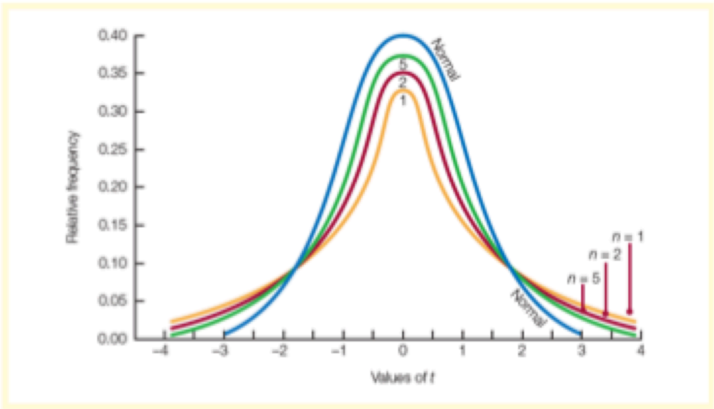
TESTING FOR *DIFFERENCES* BETWEEN VARIABLES

Independent samples t-test

Used to test a hypothesis stating that the mean scores on some interval/ratio interval (e.g. frequency of purchase) will be significantly different for **2 groups** (e.g. men vs. women)

It is based on a normal distribution with a mean value t = 0. Curve steepness is impacted by n (the *degrees of freedom*).

Exhibit 12.10 The t-distribution for various degrees of freedom



Null hypothesis - the two means are equal (e.g. male mean = women mean)
Alternative hypothesis - the two means are not equal

You should **reject the null hypothesis** when the difference between the observed t-value and the expected/critical t-value is….

Paired samples t-test

Same as independent, but the variables are **linked** in some way (e.g. husband and wife, same people at 2 different times, etc)

Analysis of variance (ANOVA)

Used to investigate the differences between **3 or more** groups on an interval variable (e.g. comparing GPA's of undergrad students, postgrad students and diploma students).

If the grouping variable (i.e. type of student) is responsible for the difference in GPA's, then the variation in GPA's between these 3 groups will be **comparatively larger** than the variation in responses *within each group*.

Uses the **F-test** to compare mean variance (larger sample variable is divided by smaller sample variance).

- F-ralue: ratio of variation explained to variation unexplained by the regression
- A **calculated F-ratio that exceeds the critical F-ratio** indicates that the results are statistically significant (null hypothesis = rejected).

Statistical vs. practical differences

Even is there is a *statistical difference,* there may not be a *practical/substantive difference*. Keep this in mind! **Statistical significance does not automatically mean important to the researcher/manager.**

- **Ecological fallacy** - assuming that the average of a group applies to everyone in the group (e.g. "Person A bought a perfume bottle, therefore Person A must be female because females buy more perfume on average than males")

TESTING FOR *ASSOCIATION* BETWEEN VARIABLES

Three ways of measuring the association between variables:

- Chi-square (two nominal variables)
- Pearson's correlation **OR** bivariate regression (two ratio variables)
- Spearman's correlation (two ordinal variables)

Chi-square test for goodness of fit

Helps us **compare** the observed frequencies with expected frequencies based on theoretical ideas about the population distribution. It allows us to test whether **two nominal variables are associated** (so this is a *nonparametric* test)**.**

You can use **two cross-tabulations** (observed and then expected) of the two variables to calculate chi-square. For example, take the question "Is brand awareness independent of respondent's sex?"

Null hypothesis = Brand awareness is independent of the respondent's sex (in other words, if 60% of sampled people are aware, then aware men = 65/100 x 0.6 = 39).

Observed	Men	Women	total
Aware	50	10	60
Unaware	15	25	40
total	65	35	100

$$\frac{(50-39)^2}{39} = 3.1$$

Expected	Men	Women	total
Aware	39	21	60
Unaware	26	14	40
total	65	35	100

Calculation	Men	Women	total
Aware	3.1	5.7	
Unaware	4.7	8.6	
Chi-Square			22.2

The calculated chi-square = 22.2, which occurs in less than 1% of occasions by chance (don't need to know this). So we reject the null hypothesis.

- Note: When you are reading a chi-square calculation from a data analysis program like SPSS, you'll be given an **'Sig Fig'** (percentage of times that this chi-square would occur). ***For all that are less than 0.05, the null hypothesis should be rejected.***