Basic Biostatistics

Section Three

Our clinical example:

- Our prior testing revealed that Drug B is better than Drug A.
  - Should I change my practices and prescribe this new drug?
- What test should be used to compare the data and aid me in answering the question?

Choosing the appropriate test

- Based on the two aspects of the data
  - Types of variables
    - Remember Day 1 with dependent and independent variables and the categories (discrete/ continuous)
  - Number of groups being compared
    - Most common questions compare two groups (dead/alive, treated/untreated, diseased/non-diseased, drug A/ drug B)
  - If the wrong test is used false conclusions can be reached

A quick review

- Discrete
  - Nominal
    - Dead/Alive, race, marital status
  - Ordinal
    - Stage of cancer, rating systems for disease i.e. heart failure or burns
- Continuous
  - Ratio
    - Age, BP, Cholesterol, Height
  - Interval
    - Quality of life, mental status exam scores, IQ

Signal to Noise

a funny statistician's term

- Standard Deviation reflects the variation of individual values around the sample mean
- Standard Error reflects the variation of many sample means from the same size samples around the true population means.
    - Standard error is always smaller than Standard Deviation.
    - Studies use multiple centers each with their own subset of data to compare to the larger sample.

Discrete/Categorical Data

- **Chi Square test** for sample size greater than 30 or no less than 5 in any combination
  - If not greater than 30, then **Fisher's Exact Test**
  - Assumes all patients are independent
  - Based on frequencies, not percents
- Results tell us if distributions are different but not the magnitude or any clinical significance
  - Group A and Group B are not the same population
  - Can not differentiate how different

Our question

Drug Group	Glycemic Control achieved	Glycemic Control not achieved
Drug A	300	80
Drug B	170	180

The Chi Square statistic for this table is 73.31 and corresponds to a p value of 0.001, available in a table. This is less than our usual 0.05 and we then reject the null hypothesis and accept that patients on Drug A achieve glycemic control at a different rate than those on Drug B, the means infer that A is superior.

For two independent groups with a continuous variable

- Student’s T-Test
  - Compares the difference in the means of two groups divided by the standard error
  - The higher the difference and the smaller the standard error the more likely the result will be significant
  - If the groups are not independent then a paired t-test is used (the groups are related in some way other than the variables in your study)
    - **Pre/ Post treatment**
    - **Eyes**
    - **Twins**

In our example

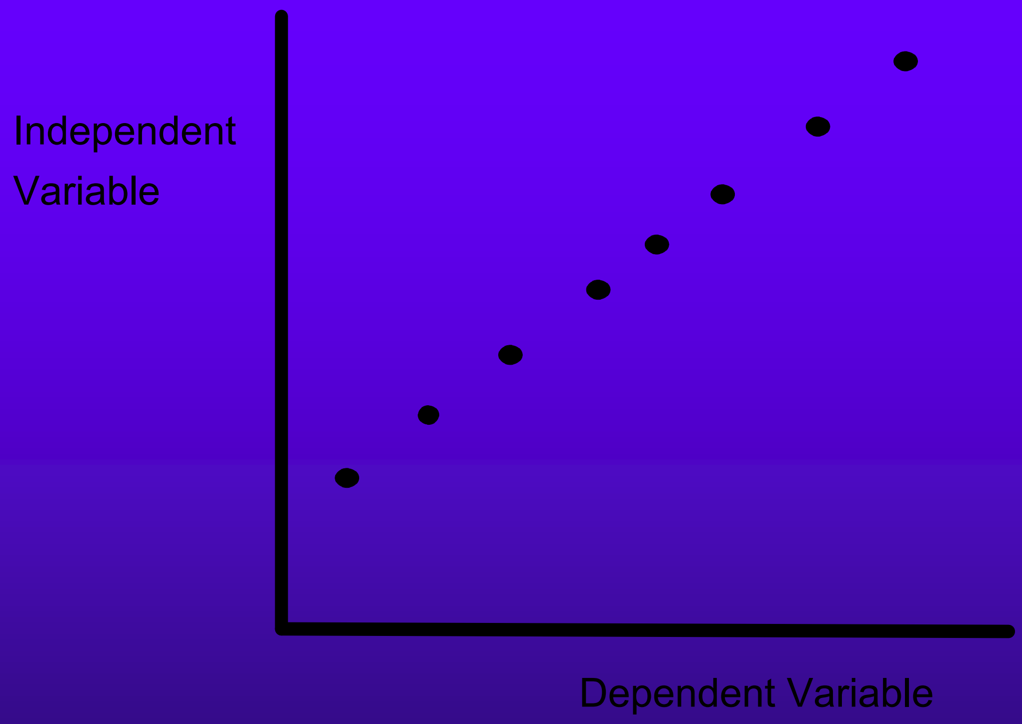
- Change the variable now Drug A and the mean blood glucose with a standard deviation
  - Compared to…
- Drug B and the mean blood glucose with a standard deviation

- We can now see the difference as a continuous number (the difference of the means divided by the standard error)

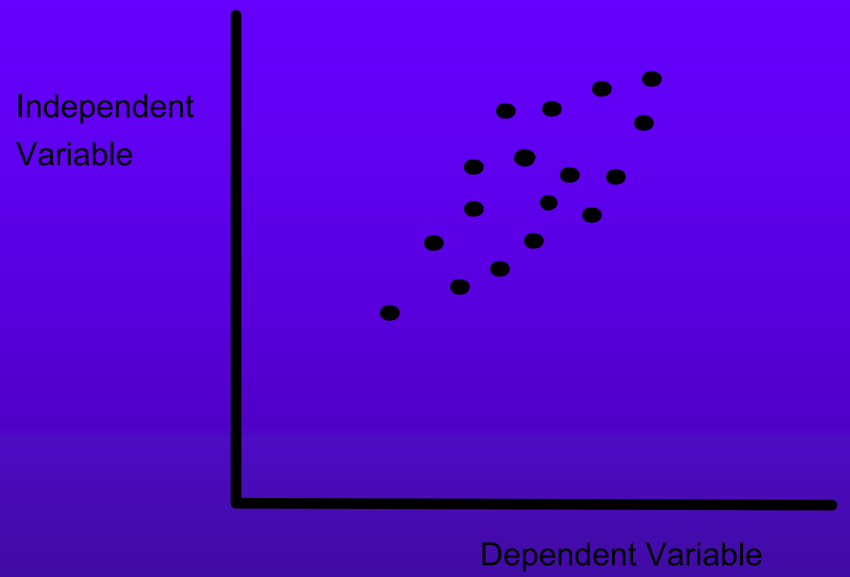
Correlation Coefficient

- Defined as r
- Describes the tightness of the data points around the best fit line
  - r = 1 is perfect positive correlation (all the dots line up in a straight line with a positive slope)
  - r = -1 is perfect negative correlation (all the dots line up in a straight line with a negative slope)
  - r > 0.7 is good, 0.5 to 0.7 is moderate and <0.5 is weak
    - If you are looking for a negative correlation then the values are negative for r

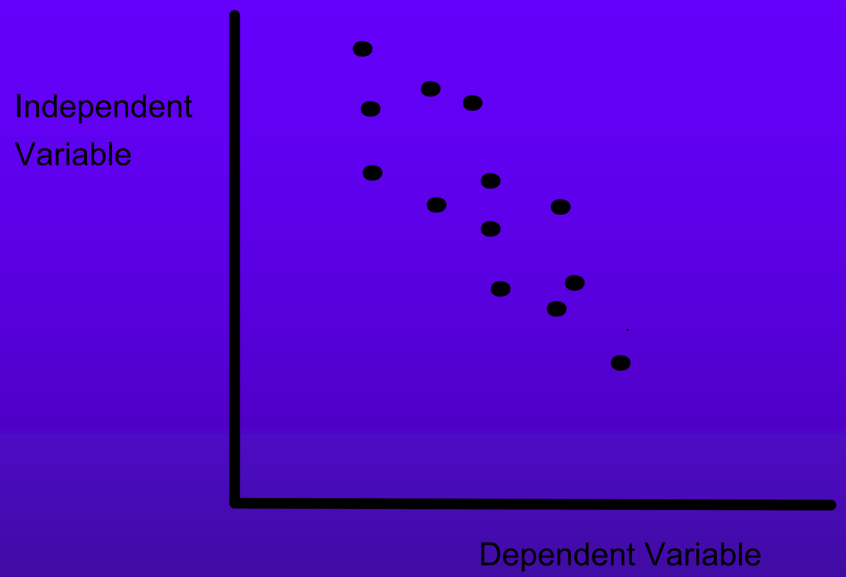
Perfect Positive Correlation



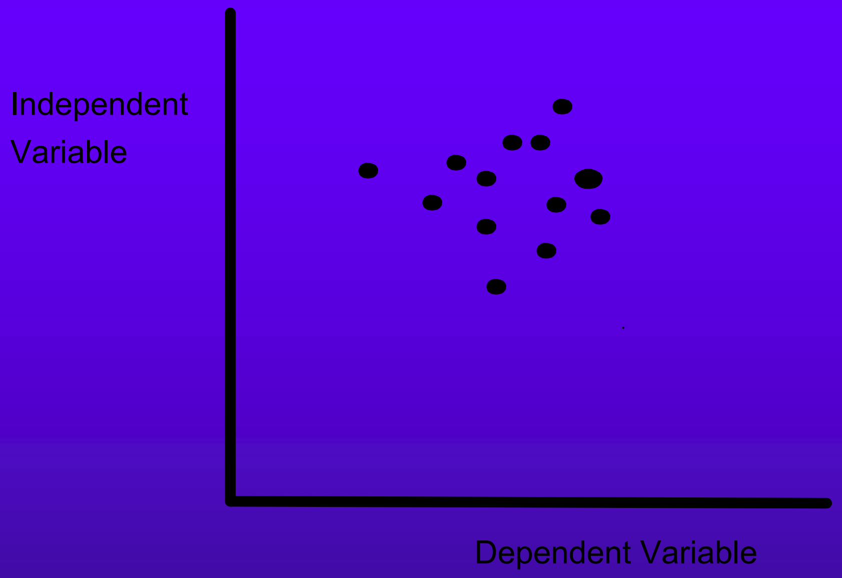
Weak Positive Correlation



Strong Negative Correlation



Nonexistent (Zero) Correlation



Pitfalls

- Correlation does not yield causation
  - Just because they are related does not mean that one causes another (coffee consumption is related to lung cancer but there is no know causal relationship)
- Correlation Coefficient squared r^2 is the amount of variance in X (the dependent variable) explained by Y (the independent variable)
  - How much of the outcome is X can be related to Y
  - Example: How much of the rate of lung cancer can be related to tobacco use (packs/day)
- The slope of the line describes the degree of the relationship
  - Back to geometry $y=bx + a$

How you use this:

- The slope (b) describes that a specific change in X that will a specific change in Y
- A p-value also exists for the slope
- $Y = bX + A$
  - b is the slope
  - b relates to r by the standard deviation
  - Each increase in packs per day causes what increase in the rate of Lung Cancer

So, finally back to our question

- In our patient with mild type II diabetes is Drug A better than Drug B and if so by how much?
  - One step closer to clinical relevance
  - Confidence Intervals

Confidence Intervals

- Another approach to p-values
- Provides a range in which we are 95% confident the true population parameter lies.
- We accept that there is a 5% chance that the real value is not in our range
- It could be comparing 2 means, odds ratios, relative risks, response rates, etc…

Confidence Intervals

- Important:
  - The interval states we are 95% sure that the real value is in our range

NOT

  - 95% of the time the value is in our range

or

  - 5% of the time our mean value is wrong

Confidence Intervals

Constructed in the formula

Sample estimate (mean, difference between means,mortality rate, odds ratio)

$\pm$ Z score for confidence

x (Standard Error)

i.e. 100 $\pm$ 5 95% CI: (95,105)

we are 95% sure that the real value is between 95 and 105

Z score and Standard Error

- The Z score is actually based on 97.5% (2.5 SD) because the 5% need to be divided on both ends of range this is 1.96

- The Standard Error is the standard deviation divided by the square root of sample size

Clinical Utility

- A confidence interval is reported for each value and if they do not overlap we can conclude with 95% confidence that they are different
- The difference can also be compared and given a confidence interval and if it doesn't cross zero you can conclude that the difference is significant and gives you a range that the difference could occur it
- An increase in power will decrease the size of the CI.

Now a new study compares

- Sugarlow to Glucobust for the treatment of Type II diabetes and reports the mean blood glucose of patients in each treatment arm
  - Sample size = 50
  - Sugarlow = 92 $\pm$ 8
  - Glucobust = 103 $\pm$ 12

Now a new study compares

- Sugarlow
  - 92 $\pm$ 8
  - 92 $\pm$1.96(8/ sqrt 50)
  - 92 $\pm$ 2.2
  - 95% CI: (89.8,94.2)
- Glucobust
  - 103 $\pm$ 12
  - 103$\pm$1.96(12/sqrt 50)
  - 103$\pm$ 3.3
  - 95% CI: (99.7,106.3)

Since the two intervals do not overlap the population differences are statistically significant.

One step more

- The difference in the values can also be compared
  - 103-92 = 11
  - Trust me the standard error is 2
    - If you want to know…. Read Studying a Study in the recommended readings.
  - 11 + 2 CI: (9,13)

Finally!

- We are 95% certain the difference is between 9 and 13
- Now for the clinician's interpretation
  - The difference could be 9
    - Worst case scenario
  - The difference could be 13
    - Best case scenario
- Is this difference clinically useful?
  - Still up to the individual clinician!
  - But you can be confident in what you decide
  - Evidence Based Medicine in a nutshell