Session 2

Quantitative Tools

- Sampling and Estimation
- Statistical Tests
- Correlation
- Robust Statistics

Sampling and Estimation

- Simple Random Sampling: Selecting a sample in such a way that each item in the population being studied has the same likelihood of being included in the sample

- Sampling Error: The difference between a sample statistic and its population value,
  - sampling error of mean = $\bar{x} - \mu$

- Sampling Distribution: a probability distribution of all possible sampling statistics computed from a set of equal size samples

Central Limit Theorem

- For a simple random samples of size n from a population with a mean μ and a finite variance σ^2, the sampling distribution of the sample mean $\bar{x}$ approaches a normal probability distribution with a mean μ and a variance $\frac{\sigma^2}{n}$ as the sample size becomes large

- Large in practice: n>=30

- If σ is unknown, use $s_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{i=n}(x_i - \bar{x})^2}{n-1}}$

Confidence Intervals

- An estimator of the population mean μ : $\bar{x} = \frac{\sum x}{n}$
- Confidence interval: range of values withinwhich the actual value willie given the probability 1 – α
  - α is the level of significance
  - 1 – α is the degree of confidence
- [$\bar{x}$ - R * SE ; $\bar{x}$ + R * SE]
  - R = Reliability factor: depends on the sampling distribution of the point of estimate and the probability that the point estimate falls in the confidence interval (1 – α)
  - SE: standard error of the point of estimate
- For a standard normal distribution R = z
  - z = 1.645 for 90% confidence intervals — α = 10%, 5% each tails
  - z = 1.96 for 95% confidence intervals — α = 5%, 2.5% each tails
  - z = 2.575 for 99 % confidence intervals — α = 1%, 0.5% each tails

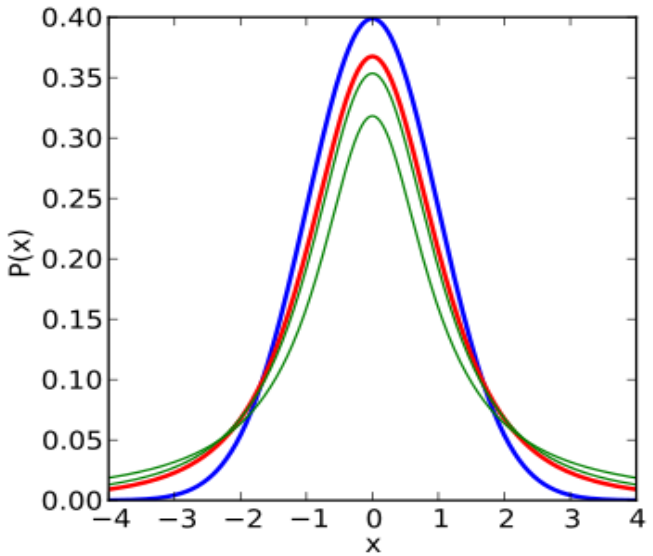
Dealing with uncertainties

- Wisdom
  - History / Experience
  - Intuitions
  - Diversifications
- Mathematics
  - Monte Carlo Simulations
  - **Statistics**

Student's t distribution

- A family of continuous probability distributions that arise when estimating the mean of a (approximately) normally distributed population in situations where the sample size is small (n<=30) and population standard deviation is unknown
- It is defined by the degrees of freedom (df)
  - df = n – 1 for sample means
- It is a symmetric bell-shaped distribution lik the normal distribution but with heavier tails.
- In Excel: function T.Inv($\alpha/2$, df)

Density of the *t*-distribution (red) for 1, 2, 3 degrees of freedom compared to the standard normal distribution (blue).



Source: en.wikipedia.org

Estimating the mean
Selecting the Appropriate Test Statistic

Distribution	Variance σ known		Variance σ unknown	
Normal	**z statistic**		**t statistic**	
Non Normal / Unknown	N small	N large	N small	N large
		z statistic		**t statistic**

- z statisitic: $[\bar{x} - z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} ; \bar{x} + z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}]$
- t statistic: $[\bar{x} - t_{\alpha/2} * \frac{s}{\sqrt{n}} ; \bar{x} + t_{\alpha/2} * \frac{s}{\sqrt{n}}]$
  - $t_{\alpha/2}$: t-distributed random variable with n-1 degrees of freedom
  - s: sample standard deviation
- The sample need to be random
  - otherwise the central limit theorem does not apply!!

Courtroom trial

- A statistical procedure is comparable to a criminal trial

- H_0: The defendant is not guilty – null hypothesis
- H_a: The Defendant is guilty – alternative hypothesis

	H_0 **is true** **Truly not guilty**	H_1 **is true** **Truly guilty**
Accept H_0 Decision: Acquittal	Right Decision	Type II Error False negative β = **Prob(Type II Error)**
Reject H_0: Conviction	Type I Error False Positive **Significance level** α **= Prob (Type I Error)**	Right Decision **Power of the test** **= 1 -** β

Source: wikipedia

Statistical Hypothesis Testing

- Is a result statistically significant?

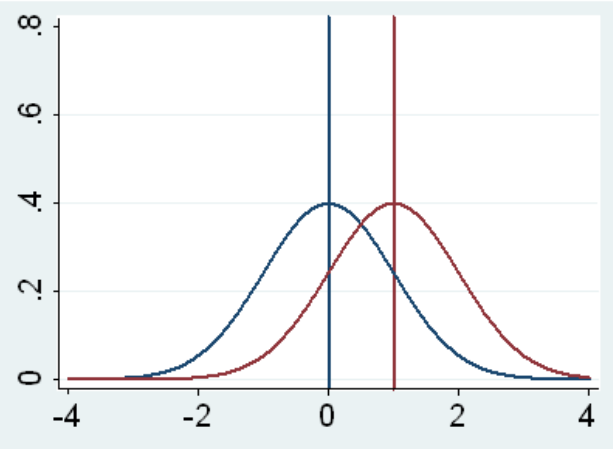
- Procedure
  - State the hypotheses
  - Set criteria for a decision
    - Select the appropriate test statistic
    - Select the level of significance
  - Collect sample from the population
  - Make a decision

Applications: A/B Tests

https://vwo.com/ab-testing/

- 2 worlds: one with the new feature, one without the new feature
- Why?
  - Impact of a new feature
  - Compare data with intuition, to better understand how users respond to certain parts of a product. You may assume certain reactions from the users

A/B Testing



- Define your goal and form your hypothesis
- Identify a control and a treatment
- Identify key metrics to measure
- Identify what data needs to be collected
- Make sure that appropriate logging is in place to collect all necessary data
- Determine how small of a difference you would like to detect
- Determine what fraction of visitors you want to be in treatment
- Run a power analysis to decide how much data you need to collect and how long you need to run the test
- Run the test for at least this long
- First time trying something new: run a A/A test to check for any systematic biases (dummy)