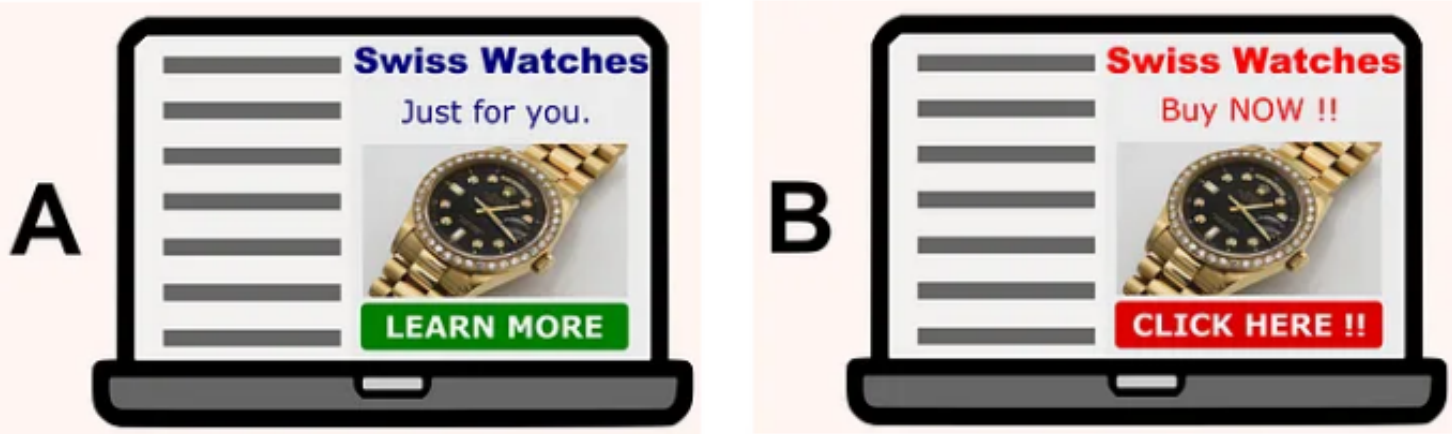
Comparing Several Samples With Applications in AB Testing and Sports

Overview

Motivation: Comparing Two Proportions

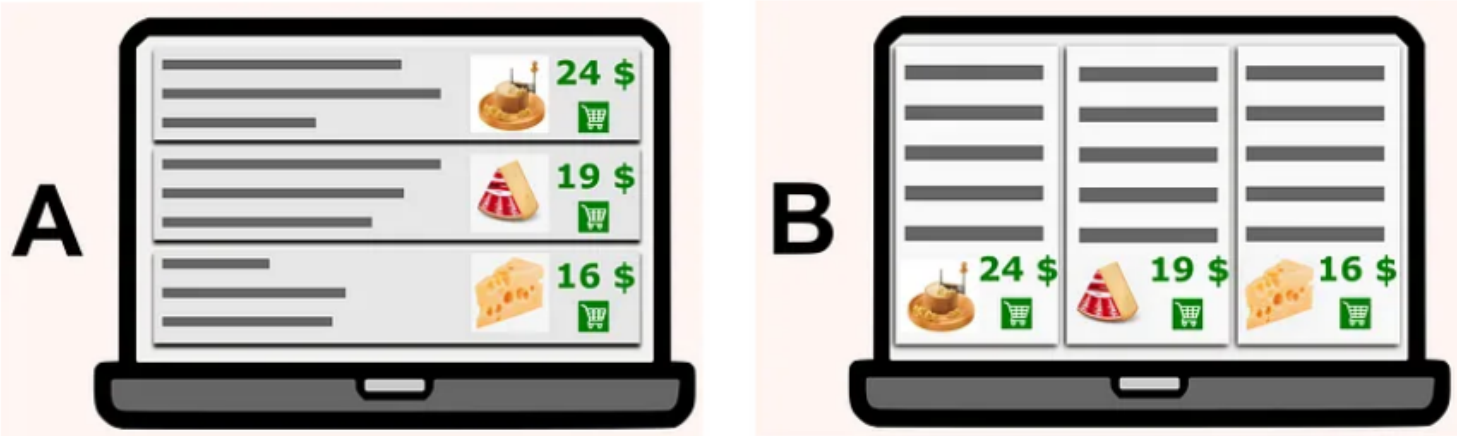
1. Click through rate: if a user is shown two adds, which one do they click on the most?



2. Conversion rate: compare the proportion of customers who follow through with a sale based on two add
3. Winning proportion between home games and away games

Motivation: Comparing Two Means

1. Average order value between two display formats



2. Average session duration between two website formats
3. Average points won between home and away games

Identifying and Estimating the Target Parameter

The unknown population parameter (e.g., mean difference or difference between proportions) that we are interested in estimating is called the target parameter

Parameter	Statistic	Key words	Data Type
$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	Mean difference	Numerical
$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	Difference between Proportions	Categorical

Comparing Two Proportions

A Confidence Interval for the Difference Between Two Proportions

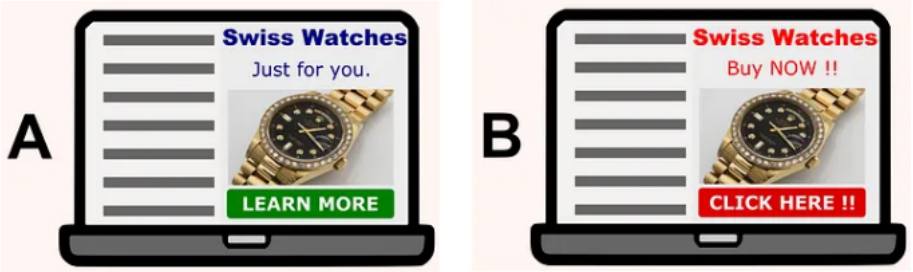
$$(\hat{p}_1 - \hat{p}_2) \pm z\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

where

- $\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$, is the **standard error** of the difference
- z is the critical value (or z-score) which depends on the confidence level

Example: Comparing Two Proportions

1. Click through rate: if a user is shown two adds, which one do they click on the most?



	A	B	tot
click	7	15	22
no click	8	4	12
tot	15	19	

2. Winning proportion between home games and away games

Comparing Two Means

A Confidence Interval for Difference Between Two Means

$$(\bar{x}_1 - \bar{x}_2) \pm t\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where

- $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, known as the **standard error**, is the estimated standard deviation of the sampling distribution of $(\bar{x}_1 - \bar{x}_2)$
- t is the critical value of the t-distribution with *adjusted* degrees of freedom [use software]

Example: Comparing Two Means

1. Average points won between home and away games

Confidence Interval Paired Data

Example

When the data is paired, you conduct one sample inference on the paired differences. Example:

- Are fathers older than mothers, on average?

ANOVA: Comparing More than Two Means

Motivation

A pharmaceutical company tested three formulations of a pain relief medicine for migraine headache sufferers. For the experiment, 27 volunteers were selected and 9 were randomly assigned to one of the three drug formulations. The subjects were told to take the drug during their next migraine episode and to report their pain on a scale of 1 = no pain to 10 = extreme pain 30 mins after taking the drug.

- Is there a difference in the mean pain level of the patients among the 3 drug formulations?
- Post hoc: Knowing that the means differ leads to the question of which groups have different means?

ANOVA F-Test

1. Check Assumptions: Random samples or random assignment? Independence within groups? Nearly normal populations? Independent groups? Equal population variances?
2. State Hypotheses:
   $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ (The population means are all equal)
   H_A : At least one population mean is different
3. Compute Test statistics: $F = \frac{MSB}{MSW}$.
4. Compute p-value. Make a decision based on a pre-specified α. [Note: if p-value$< \alpha$ reject H_0, else do not reject.]
5. Conclude

Software Output for ANOVA

Source of Variation	DF	SS	Mean Square	F	P
Between (Treatment)	$k-1$	SSB	$MSB = \frac{SSB}{k-1}$	$F = \frac{MSB}{MSW}$	
Within (Error)	$N-k$	SSW	$MSW = \frac{SSW}{N-k}$		
Total	$N-1$	SST			