A/B testing (aka controlled experiments):

Hypothesis Testing:

H0 = opposite of your assumption (people like coffee as much as tea)

H1= your assumption (people prefer coffee over tea)

The idea behind hypothesis testing is that we can never really know if an assumption is true from data, but we can show how unlikely it would be to get the data we currently have if our assumption was false.

In simple words- from sample you cant assume something about the entire population. So we try to prove how the opposite of our assumption is not possible with given data.

<https://www.youtube.com/watch?v=X8u6kr4fxXc>

**Control group**: existing feature

**Treatment group:** new feature

Estimation of sample size:

<https://www.youtube.com/watch?v=JEAsoUrX6KQ>

Sample size= 16\* (population variance)/ (delta)2

delta= difference between mean of treatment and control population sample 🡪 we do not know the value of delta before experiments, so we use minimum detectable effect

Sample size larger if variance is high

Sample size is small in delta is small

|  |  |  |
| --- | --- | --- |
| Hypothesis testing | 1. One-tailed and two-tailed tests  2. ANOVA 3. z- test 4. t- test | Null Hypothesis:   H0: μ=μ0 (two-sided), μ≥μ0, μ≤μ0 (one-sided) Alternative Hypothesis:   Ha : μ!=μ0 (two-sided), μ<μ0 , μ>μ0(one-sided) |

|  |  |
| --- | --- |
| Z-Test | When sample size is large compares mean (u) of the sample to some value Uses population standard deviation (σ) |
| |  | | --- | | **Description**: Tests if the mean μ is equal/less than/ greater than μ0 **Statistic**: μ (mean) **Distribution**: N(μ,square of (σ)) (normal)  z =  (s=σ if known) | |
| T- test | When Sample size is small Uses sample's standard deviation (s) |
| |  | | --- | | **Description**: Tests if the mean μ is equal/less than/ greater than μ0 for small sample (n<=30) **Statistic**: μ (mean) **Distribution**: T(n-1)(t)  T=   s= σ if known | |
| |  | | --- | | Independent Two-sample t-test With **Equal Variances and Sample Sizes:  Description:** Tests if the mean of one sample (μ1) is different/less than/more than the mean of another sample (μ2) when the samples have equal sample size, variance, and are independent **Statistic**: μ1−μ2 (mean) **Distribution:** T(2n−2) (t) | |
| |  | | --- | | Independent Two-sample t-test With **Equal Variances and Unequal Sample Sizes  Description**: Tests if the mean of one sample (μ1) with n1 data points is different/less than/more than the mean of another sample (μ2) with n2 data points when the samples have equal variances and are independent **Statistic**: μ1−μ2(mean) **Distribution**: T(n1 + n2−2) (t) | |
| |  | | --- | | Dependent Two-sample t-test  **Description**: Tests if the mean of one sample (μ1) with n1 data points is different/less than/more than the mean of another sample (μ2) with n2 data points when the samples have unequal variances and are independent **Statistic**: μ1−μ2 (mean) **Distribution**: T(df) (t) where    **Test Statistic**: | |
| Paired (Dependent) t-test **Description**: Tests if the ***difference of the same sample before and after some treatment*** is equal/less than/more than some number μ0 **Statistic**: μΔ (mean of differences between pairs) **Distribution**: T(n−1) (t) |
| Proportion Testing | Works for categorical data i.e. Binomial Distribution e.g: (% voters that vote for candidate A, % of people who prefer coffee to tea, % of images correctly classified by a machine learning model) |
|  | **One sample Test Proportion**  **Description**: Tests if a proportion of a sample (p) is equal to or less than or more than some p0 ∈(0,1) **Statistic**: p (proportion) **Distribution**: B(n,p0) (Binomial)  z = (p−p0)/ sqrt(p0(1-p0)/n)  p= no of success observations/n |
|  | **Two Sample Test of Proportions**  **Description**: Tests if a proportion of a sample (p1) is equal to/less than/ more than a proportion of another sample (p2) **Statistic**: p1 −p2 (difference of proportion) **Distribution**: Literally B(n1 ,p1)−B(n2,p2) (Difference of Binomials), but practically N(p1−p2,v) (Normal Approximation)  z= (p1-p2)/ sqrt(v(1-v)(1/n1 + 1/n2))  v= Number of oberservations in the success class across BOTH samples / (n1+n2) |