# Video 1: non-pamaetric testing the permutation test.

wbg231

December 2022

## 1 introduction

- vedio link

### 1.1 hyptoshsi testing

- steps of hypothsis testing framework
  1. chose a conjectture ie allternative hypothsis $H_1$
  2. chose a null ie $H_0$
  3. chose a test stat (should be large if there is a lot of evedince against the null)
  4. decide on a signefgence level
  5. gater data and compute test stat
  6. compute the p value
  7. reject the null if p-value $\leq \alpha$

### 1.2 p-value

- the p-value is the probability of observing a larger or equal test stat under the null hypothesis
- the issue with that is we need to know how the test stat is distributed under the null
- what if we do not know this?

### 1.3 price of burgers example

- suppose we have a conjecture that burgers in NYC are more expensive than those in madrid
- null is that the price of burgers in both cities has the same distribution
- the test stat is average price of burger in nyc - average price of burger in madrid
- this would be high if the alternative holds and low otherwise
- suppose we see two places in nyc that charge 16 and 18 dollars each
- and two places in madrid charge 13 dollars each
- $d_{data} = m(ny) - m(\text{Madrid}) = 4$
- is this evedince against the null? we dont know with out a p-value
- but we do not know how our test stat is distributed under the null

### 1.4 key idea

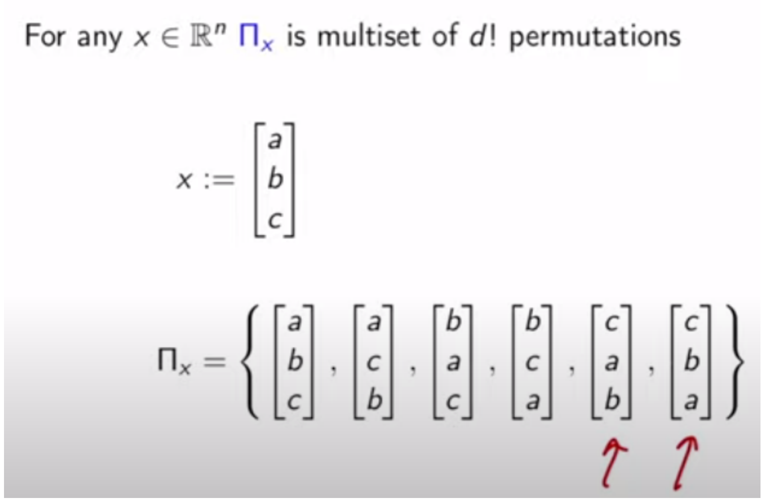
- if the price distribution is the same in both cities where the burgers came from should not matter
- that is we should be able to switch around the labels and if it holds we would have been equally likely to observe that data as well
- so we can look at all permutations of the labels and there corresponding test statistics
- here are all possible realizations of the test stat
- 

| NY | NY | M | M | t |
|---|---|---|---|---|
| 13 | 13 | 16 | 18 | -4 |
| 13 | 13 | 18 | 16 | -4 |
| 13 | 16 | 13 | 18 | -1 |
| 13 | 16 | 18 | 13 | -1 |
| 13 | 18 | 13 | 16 | 1 |
| 13 | 18 | 16 | 13 | 1 |
| 13 | 13 | 16 | 18 | -4 |
| 13 | 13 | 18 | 16 | -4 |
| 13 | 16 | 13 | 18 | -1 |
| 13 | 16 | 18 | 13 | -1 |
| 13 | 18 | 13 | 16 | 1 |
| 13 | 18 | 16 | 13 | 1 |

| NY | NY | M | M | t |
|---|---|---|---|---|
| 16 | 13 | 13 | 18 | -1 |
| 16 | 13 | 18 | 13 | -1 |
| 16 | 13 | 13 | 18 | -1 |
| 16 | 13 | 18 | 13 | -1 |
| 16 | 18 | 13 | 13 | 4 |
| 16 | 18 | 13 | 13 | 4 |
| 18 | 13 | 16 | 13 | 1 |
| 18 | 13 | 13 | 16 | 1 |
| 18 | 16 | 13 | 13 | 4 |
| 18 | 16 | 13 | 13 | 4 |
| 18 | 13 | 13 | 16 | 1 |
| 18 | 13 | 16 | 13 | 1 |

- so now we want to know how unlikely was the data we saw.
- that is how many of these permutations have a test stat larger or equal to that we actual observed in the data (4)? it was like 16.7%
- this is not quite the same

### 1.5 multiset of permutations

- for any $x \in \mathbb{R}^d \Pi_x$ is a multiset of $d!$ perumations
- 

- so noice here that the last two elements of the multiset are the same, so it is not a propper set

### 1.6 p-value of permuation test

- if our observed data is $x_{data} \in \mathbb{R}^d$
- and we have observed test stat $t_{data} := T(x_{data})$
- and we mode the distribution of the data under the null ypothsis as a random vector $\tilde{x}_{null}$
- then our test stat unde rthe null we call $\tilde{t}_{null} = T(\tilde{x}_{null})$
- so our p value function in this case is $pv(t) := P(\tilde{t}_{null} \geq t | \tilde{x}_{null} \in \Pi_{x_{data}})$
- so that is the p -value is teh probaility that t under the null is greater than that value given teh data we observed is a a permuation of the observed data
- in the typical p-value we do not condtion on this, but this is required for the permuation test

## 1.7 exchangability

- the enteries of $\tilde{x}$ are exchangable if permuting them does not change the distrobution of $\tilde{x}$
- $\Pi_x$ the multset of permautions of x
- the enteries fo a discrete random vector $\tilde{X}$ are exchangable if

$$\forall v \in \Pi_x, P_{\tilde{x}}(x) = P_{\tilde{x}}(v)$$

- so that means changing the order of ellements does not change the distrobution of the vector
- the same holds for continous rv

## 1.8 idd rvs

- exhangability will hold for iid rvs
- we the entries are indepdent then the joint is the product of the marginals then if they are identically distributed it is the asme as the marginal of eahc multiples toghter. then if we look at any vecotr in the multset so exchangability holds
- if $\tilde{x}_1 ... \tilde{x}_d$ are iid
- then $f_{\tilde{x}}(X) = \Pi_{i=1}^{d} f_{\tilde{x}_i}(x_i) = \Pi_{i=1}^{d} f_{marg}(x_i) = \Pi_{i=1}^{d} f_{marg}(v_i) = f_{\tilde{x}}(x), \forall v \in \Pi_x$

## 1.9 consequnces of exhangability

- if $\tilde{x}_{null}$ is exhangeable then the p values are unfiorm that is
- $P(\tilde{x}_{null} = v | \tilde{x}_{null} \in \Pi_{x_{data}}) = \frac{1}{n!}, \forall v \in \Pi_x$
- this holds pretty directly from the defention of exhangability
- furhter then we know the sum of the condtionals of each value in the multiset add up to 1
- once again this holds pretty directly from the defention
- the real point is that the liklhood of seing any perumation is equal

## 1.10 nonparametric p value function

- this is important because it allows us to comptue the distribution of our p values even when we do not have acess to the distribution of the test stat under the null
- bascailly our p value of an event is the number of times that events occours over the total number of events (N!)
- more specfically we can write out our p value function as $pv(t) = P(T(\tilde{x}_{null}) \geq t|\tilde{x}_{null} \in \Pi_{x_{data}}) = P(\cup_{v\in\Pi_{x_{data}}:T(v)\geq t}|\tilde{x}_{null} \in \Pi_{x_{data}}) = \Sigma_{v\in\Pi_{x_{data}}:T(v)\geq t}P(\tilde{x}_{null} = v|\tilde{x}_{null} \in \Pi_{x_{data}}) = \frac{\Sigma_{v\in\Pi_{x_{data}}}\mathbb{I}(T(v)\geq t)}{n!}$
- we are more or less saying if the enteries are exchangable under the null than we would be equally likely to see any outcome, and to comptute the condtional prob the test stat is greater than what we saw than we kust need teh number of cases where the test stat is greater than or equal to what we saw over the liklyhood of seing it whcih is 1 over the number of permutations

## 1.11 false postive

- the point of all of this is that this re-defined p -value still allows us to garuntee that $P(\text{false postive} \leq \alpha)$
- this holds from a really close argument to what works with the regular p -value
- our new defention of p-value using this condtional arugment still lets us maintain the liklyhood of a false postive being upper boudned by the signefgence level

## 1.12 free throws examples

- we are doing the free throw example that we have done in the past few vedios
- so assume we have the same set up
- unde the null the free throws are iid
- in our data he shoots 44 free throws at home and 41 away

- so under the permuation test we can write our data as $\tilde{x}_{data} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ ... \\ 1 \end{pmatrix}$ which is of length (85) ie the number of shots at home pluess the number of shots away and where the liklhood of any permuation is equally

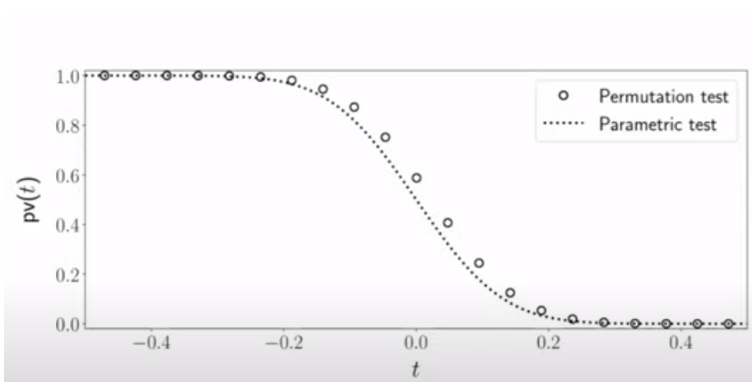
- so our test stat can be expressed as $T(v) = \frac{1}{44}\Sigma_{i=1}^{44} v[i] - \frac{1}{41}\Sigma_{i=45}^{85} v[i]$ where i is any random permuation of our input vector

- then our p value is $pv(t) = \frac{\Sigma_{v\in\Pi_{x_{data}}} \mathbb{I}(T(v)\geq t)}{n!}$

- the issue with this is that there are 85 data points and 85! is way to high to compute

- so we an use montecarlo simulation to take some random subset of the multiset to aproimate this (that is the jack knife)

### 1.13 parametric two sample vs nonparametric permuation test for free throws example

- 


- here we have plotted the p-value function for both tests

- they are super similar

- but the permuation test is a lot more simple

- this is a numerical aproximation that aproximates an analytical solution very well

- the nice thing about this test is that it is easy and works pretty well in a lot of cases

- after this he doss an example with the median i think it is worth watching if you are intrested i am not going to write it down tho.