**Scenario**: Let us consider that we have 1 Central Node(C.N) and clients ∈ {1,2 … p} with distributions D1, D2 .. Dp.

**Question**: Let’s say we consider client i, whose distribution is Di. Can we utilize the information from the other remaining clients, to update the weights of the model of client i, so that we can get better performance?

**Observation**: A clear observation is that we can only use information from the other clients if there is dependency of client i, with client j. This means, the distribution of client i, should be similar to client j i.e Di~Dj. The dependency between clients can be mathematically calculated by 2 methods: 1. K.L Divergence between distributions, 2. Discrepancy between clients.

We will be focusing on the 2nd point, i.e finding out the discrepancy between clients. Let us take a 2 client scenario. We define the discrepancy in this case as→

Let us define the model of the first client to be N1 and the data w.r.t the first client to be X1. Given this scenario, let L1 be the loss incurred in this case, w.r.t client 1. Let the data w.r.t the second client be X2. Now, let us define L2 to be the loss incurred when we use the same model as client 1, for client 2.

We define **discrepancy** as:

d = sup**w** L2(**w**) - L1(**w**)

**New Problem:** But, in the given case, we are learning in an online fashion such that we are receiving the data at client 1 and 2, one at a time. So, we wouldn’t be able to calculate the actual value of the discrepancy but only an estimate $\hat{d}$. Given each new sample, we would have to update our estimate $\hat{d}$.

**Discrepancy Estimate**: Let us denote the discrepancy at time ‘t’ between client 1 and 2, as d12,t . We need to estimate d12,t+1 based on the previous estimate. Let Lt denote the loss at time ‘t’ and Lt+1 at time ‘t+1’. We will know that our estimate of discrepancy d, is accurate when |d12,t+1-d12,t| < є.

Exact discrepancy Calculation: ?? ( Need help in this)

Let us assume that we know the discrepancy $\hat{d}$. Let us say that 𝛼1 and 𝛼2 denote the dependency of clients 1 and 2 respectively, on the weight update w.r.t client 1. We need to change 𝛼1 and 𝛼2 iteratively, so as to get a better model w.r.t client 1( as well as 2). So, we follow the specific statistic:

**if** $\hat{d}$>ꞇ // where ꞇ is some threshold(hyperparameter)

{ 𝛼1 + β1𝛼1 ; 𝛼2 - β2𝛼2 p, where β1𝛼1 = β2𝛼2

{ 𝛼1 ; 𝛼2 1-p

**else**

{ 𝛼1 ~ 𝛼2 = 0.5