Lecture 7 - 07-04-2020

Bounding statistical risk of a predictor
Design a learning algorithm that predict with small statistical risk

$$(D, \ell) \qquad \ell_d(h) = \mathbb{E}\left[\, \ell(y), h(x) \,\right]$$

were D is unknown

$$\ell(y, \hat{y}) \in [0, 1] \quad \forall y, \hat{y} \in Y$$

We cannot compute statistical risk of all predictor.
We assume statistical loss is bounded so between 0 and 1. Not true for all losses (like logarithmic).
Before design a learning algorithm with lowest risk, How can we estimate risk?
We can use test error $\rightarrow$ way to measure performances of a predictor h. We want to link test error and risk.
Test set $S' = \{(x'_1, y'_1)...(x'_n, y'_n)\}$ is a random sample from D
How can we use this assumption?
Go back to the definition of test error

Sample mean (IT: Media campionaria)

$$\hat{\ell}_s(h) = \frac{1}{n} \cdot \sum_{t=1}^{n} \ell(\hat{y}_t, h(x'_t))$$

i can look at this as a random variable $\ell(y'_t, h(x'_t))$

$$\mathbb{E}\left[\, \ell(y'_t, h(x'_t))\right] = \ell_D(h) \longrightarrow \mathit{risk}$$