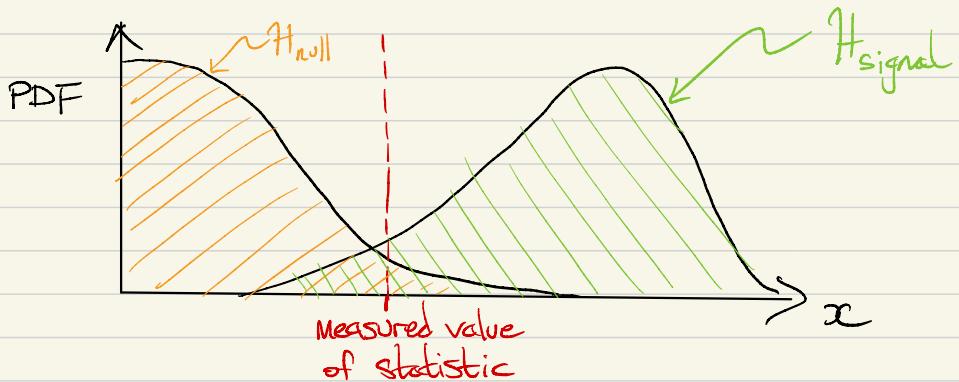


Lecture 7

Hypothesis Testing



➡ Is our experimental outcome more consistent with the **signal model**, or the **noise model**?



p-value

= probability of getting a more extreme value than measured under the null hypothesis.

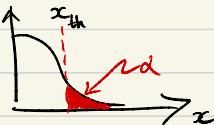
$$\stackrel{\text{def}}{=} p = \int_{x_i}^{\infty} h_0(x) dx = 1 - CDF(x_i)$$

--- Typically choose a threshold level " α " such that we can reject null if $p < \alpha$.

NOTE: doesn't mean we accept a particular signal!

TYPE I Errors

- False positives
- FP probability controlled by threshold α



TYPE II Errors

- False negatives
- FN probability controlled by variable β



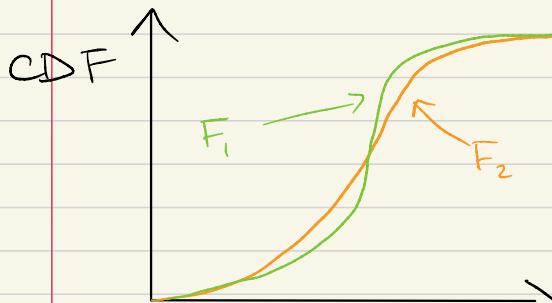
Comparing Distributions

Q

what is the probability that 2 batches of data come from the same distribution?

what is the probability that some samples were drawn from a known distribution?

e.g. Kolmogorov-Smirnov (KS) Test



$$D = \max |F_1(x_i) - F_2(x_i)|$$

for fixed FP " α " what is the threshold value of D ?

$$D_{KS} = \left[-\ln(\alpha/2) / 2n_e \right]^{1/2}$$

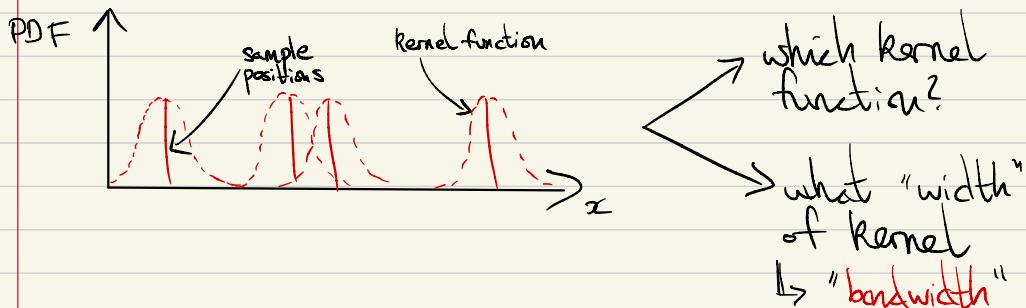
- Non-parametric modeling of histograms
 ↗ we have some data but don't know the PDF

(i) Histogram — how do you choose bin-width?
 ↗ Scott's Rule
 ↗ Freedman-Diaconis Rule

$$\Delta_{\text{bin}} \approx \frac{2.7 \sigma_0}{N^{1/3}}$$

(ii) Kernel Density Estimation

→ instead of binning datapoints, replace all with a kernel function ... smooths distribution



Kernel choices = Tophat, Gaussian, Epanechnikov

Bandwidth = (variants of) Freedman-Diaconis, cross-validation, Sheather-Jones,
 (more later)