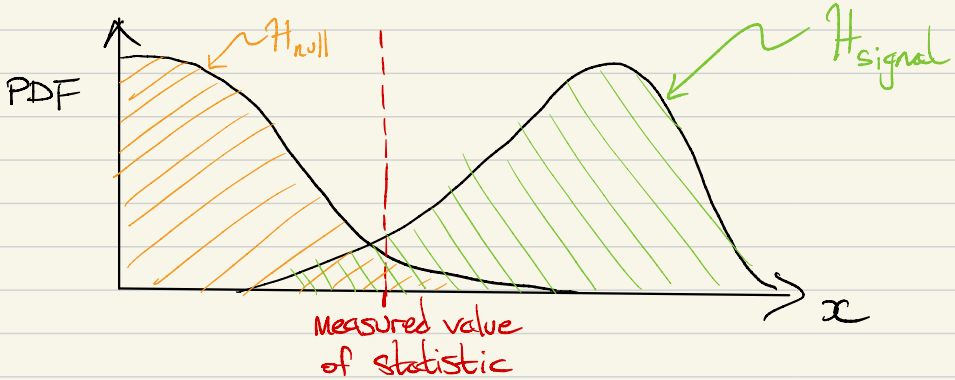
Lecture 7

02/09/2022

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.

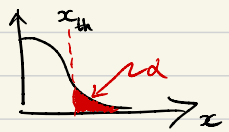
i.e. $$p = \int_{x_i}^{\infty} h_0(x)\,dx = 1 - CDF(x_i)$$

... TYPICALLY choose a threshold level "α" such that we can reject null if $p \leq \alpha$.

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

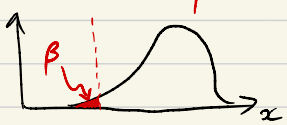
Type I Errors

- 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**

$$D = \max \left| F_1(x_1) - F_2(x_2) \right|$$

... 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 & histograms

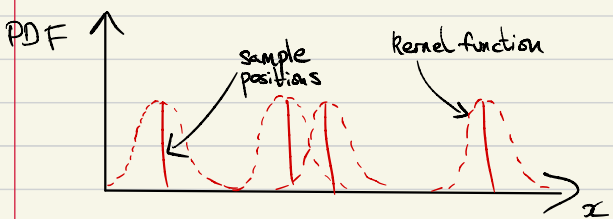
⇒ we have some data but don't know the PDF

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

$$\Delta_{bin} \simeq \frac{2.7\sigma_G}{N^{1/3}}$$

(ii) Kernel Density Estimation

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



- which kernel function?
- what "width" of kernel
  ↳ "bandwidth"

Kernel choices = Tophat, Gaussian, Epanechnikov

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