

Concentration Inequalities, Oracles, and Applications

A. DasGupta

Outline: Week 3

- 1) Uniform convergence of relative frequencies in multinomial experiments; generalization to general Empirical processes on the real line;
- 2) DKW inequality on \mathcal{R} with best constant; Glivenko- Cantelli theorem;
- 3) Central limit theorem for the Empirical process on \mathcal{R} ; Kolmogorov-Smirnov test for goodness of fit as a corollary; percentiles of the Kolmogorov-Smirnov statistics; wrong use of the Kolmogorov-Smirnov test in inappropriate problems; corrections needed;
- 4) Empirical measures in $\mathcal{R}^d, 1 < d < \infty$; question of uniform convergence of the empirical to the true; introduction of shattering coefficients; concept of VC classes of sets; Glivenko-Cantelli theorem in $\mathcal{R}^d, 1 < d < \infty$ for VC classes; DKW inequality in $\mathcal{R}^d, 1 < d < \infty$;
- 5) Concrete examples of VC classes;
- 6) Uniform central limit theorems in $\mathcal{R}^d, 1 < d < \infty$; application to goodness of fit testing in $\mathcal{R}^d, 1 < d < \infty$;
- 7) Characterizing the limiting distribution of the goodness of fit statistics; computing the cutoff values; the role of the bootstrap.