

Statistical Methods of Data Analysis

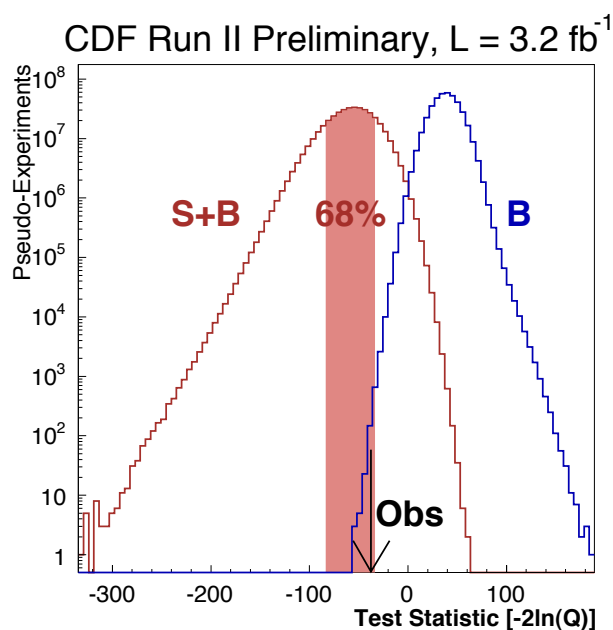
Ulrich Husemann

Humboldt-Universität zu Berlin
Winter Semester 2010/2011

Single Top Observation



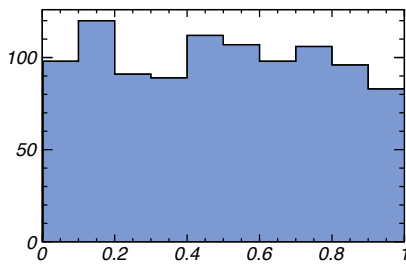
- Test statistic:
likelihood ratio
- Hypotheses:
 - S+B: single top production and background processes
 - B: only background processes
- p-value for null hypothesis B:
 $3.1 \times 10^{-7} \rightarrow 5.0\sigma$ observation



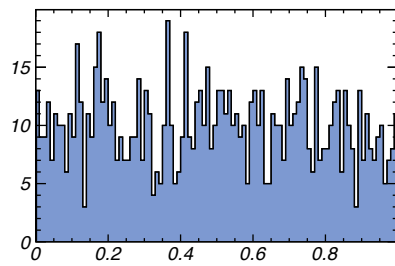
Pearson's χ^2 Test



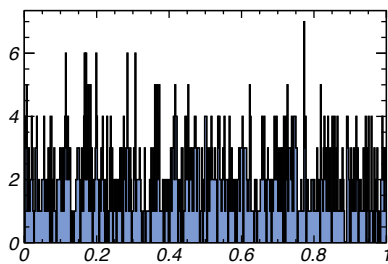
Random Number χ^2 Test (10 Bins)



Random Number χ^2 Test (100 Bins)



Random Number χ^2 Test (500 Bins)

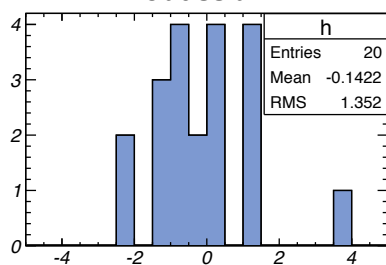


- 1000 random numbers
→ uniform distribution?
- p-values for this example
 - 10 bins: 0.25
 - 100 bins: 0.30
 - 500 bins: 0.85

Kolmogorov-Smirnov Test

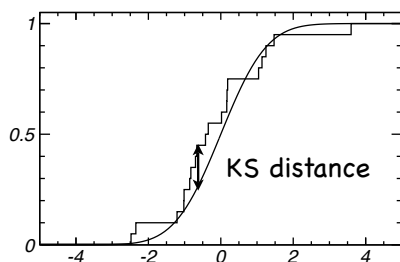


Gaussian



- 20 random numbers from Gaussian PDF
- Comparison of Gaussian CDF and empirical CDF
- KS distance = 0.196
- p-value: 0.425

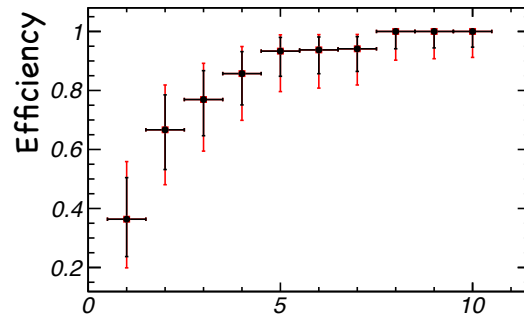
KS: CDFs



Efficiency: CIs



- Variance of binomial distribution = 0 for efficiencies of 0, 1
- Solution: construct 68.3% confidence level intervals
→ asymmetric error bars



- Solutions in ROOT 5.28
 - Black: TGraphAsymmErrors::BayesDivide()
→ Bayesian CI (default in ROOT 5.26)
 - Red: TGraphAsymmErrors::Divide()
→ Clopper-Pearson CI