An Idea for Photometric Validation of Star/Galaxy Separation

Alex Drlica-Wagner

January 22, 2014
SVA1 Analysis Meeting

Motivation

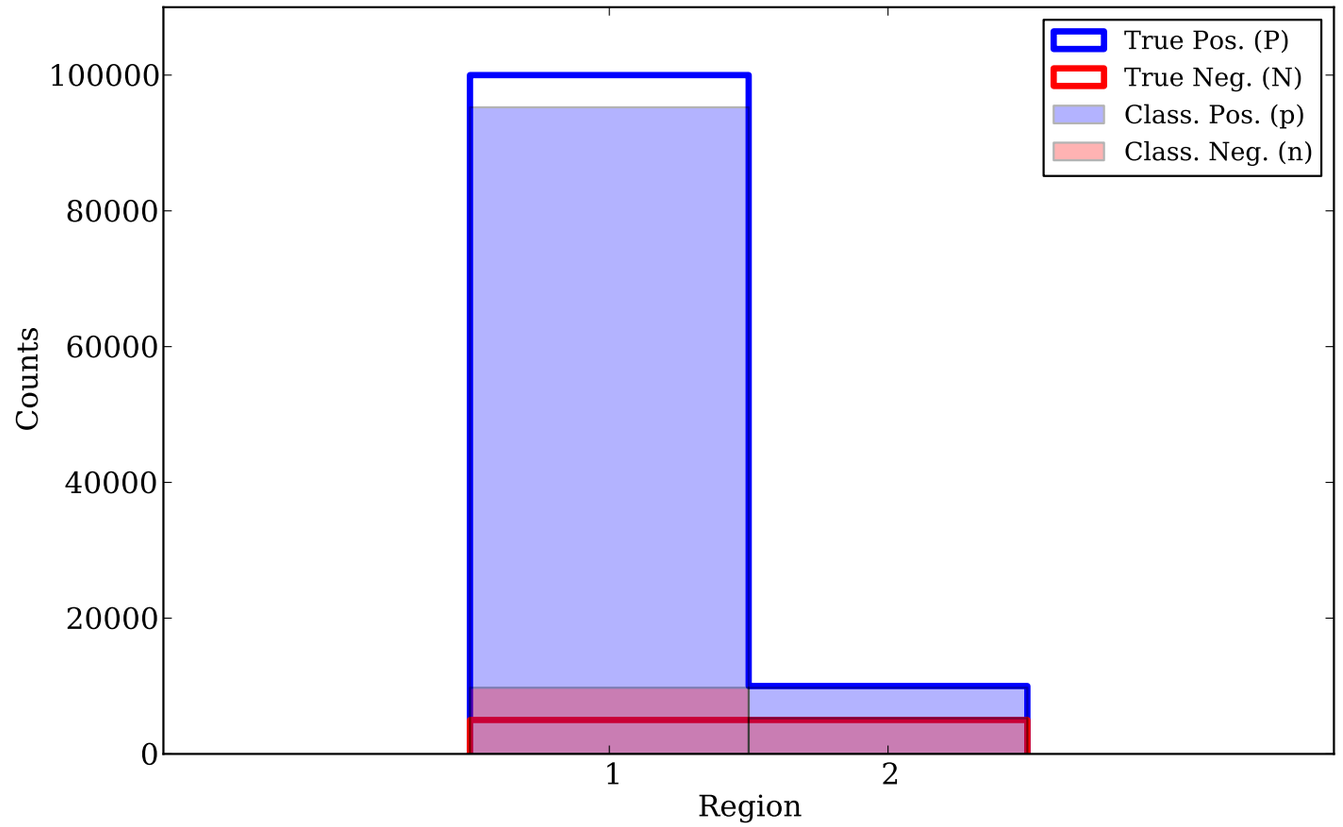
- **Training and test samples require additional observations (HST, spectroscopic, etc.), and it would nice to have a way to validate (at least roughly) the star/galaxy separation using DES data alone.**

- **The idea is to use "regions" of DES data where the distributions of stars and galaxies are distinct.**

- **Start with spatial over densities:**
  - **Galaxy clusters**
  - **Resolved stellar clusters**
- **Could image using color or magnitude information**
- **This method should be thought of as a rough cross-check, not as the final word on completeness**

Toy Example

- **Assumptions:**
  - **Positive (signal) objects dependent on region**
  - **Distribution of negative (background) objects independent of region**
  - **Performance of the classifier is the independent of region**
- **Leakage from the positive distribution into the negative distribution will depend on bin**
- **The "True Positive Rate (TPR)" (aka, completeness) can be analytically derived from the observed classification**



$$\mathrm{TPR} = \frac{\Delta p}{\Delta n + \Delta p}$$

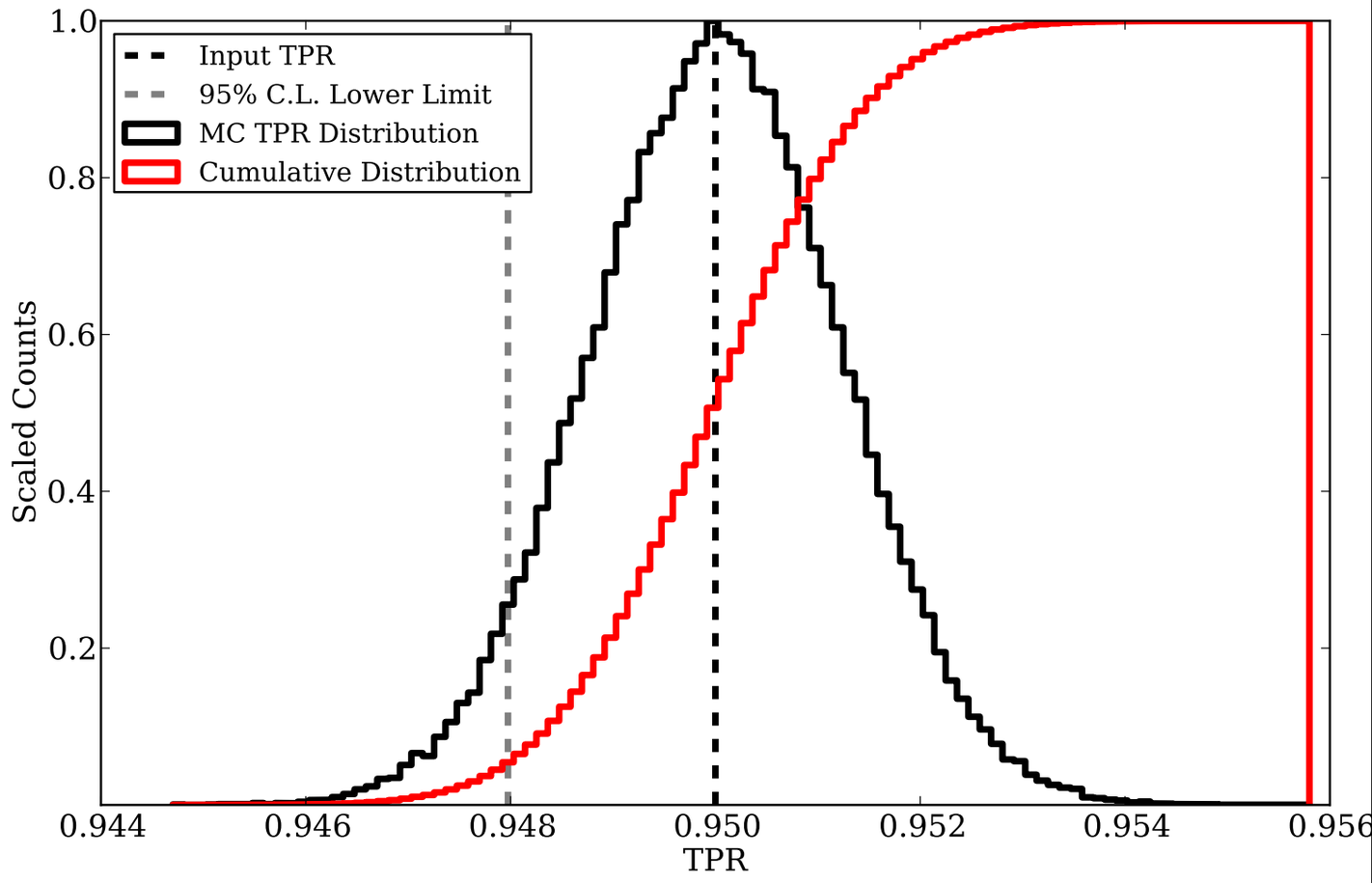
where

$$\Delta p = p_1 - p_2$$

$$\Delta n = n_1 - n_2$$

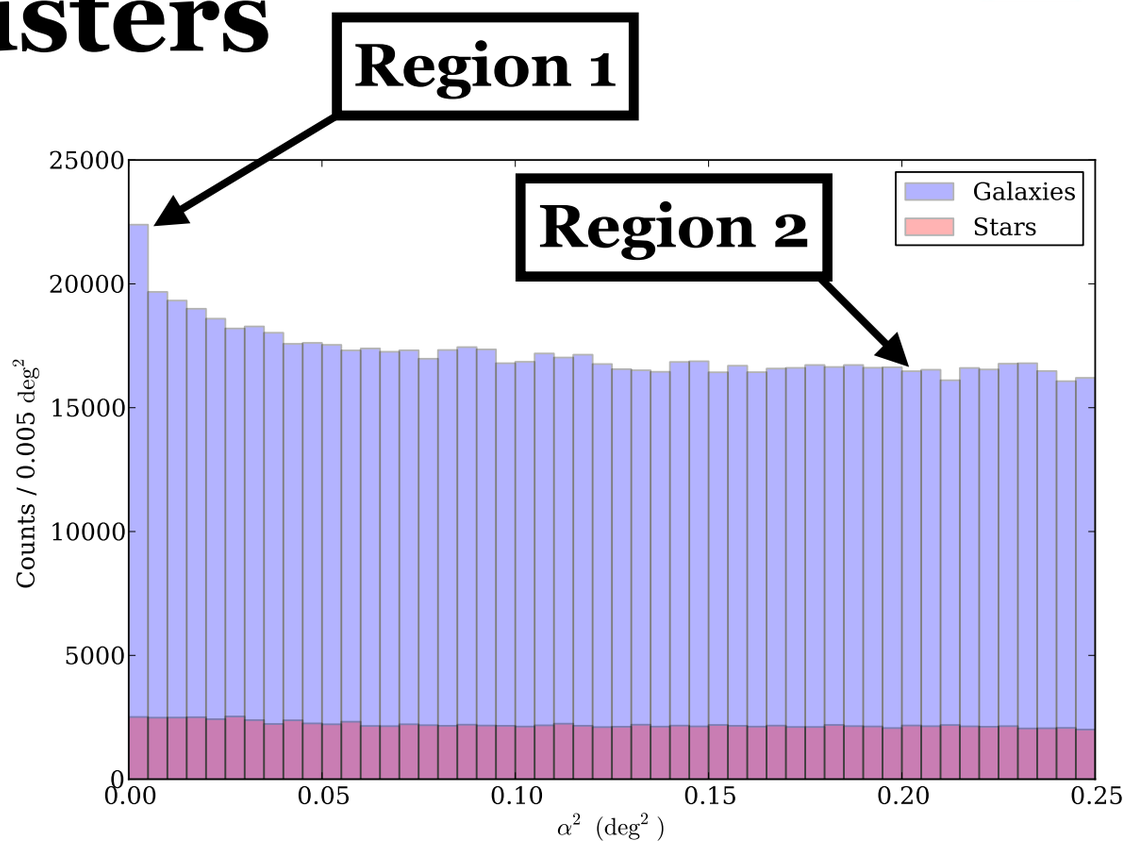
Statistical Uncertainty

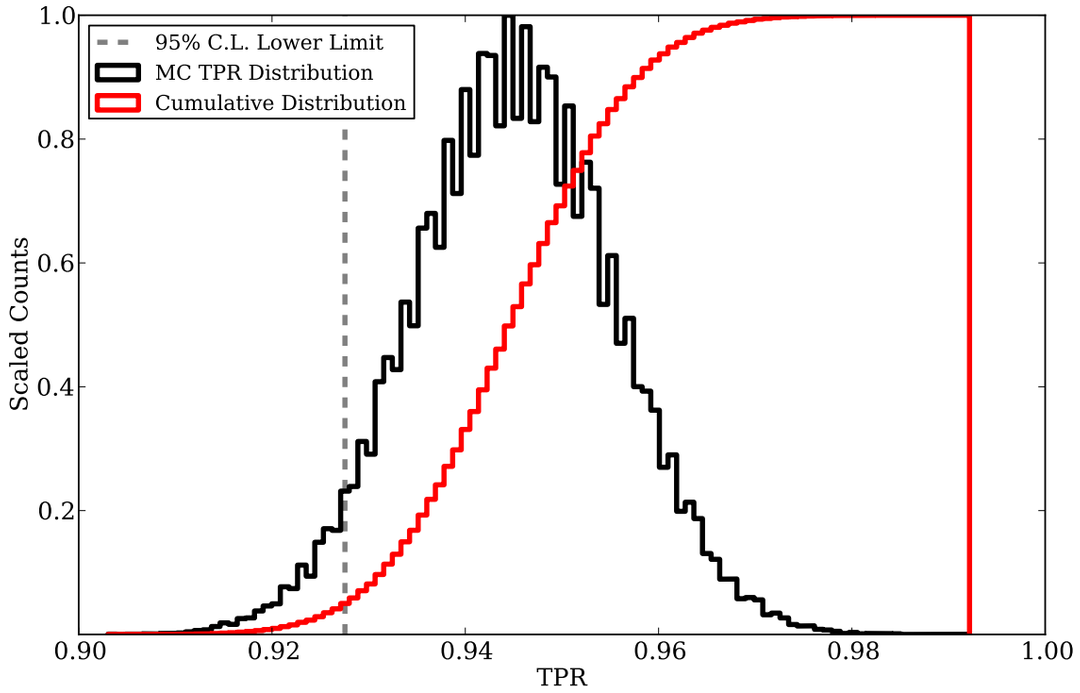
- The classification of objects in each region will be subject to statistical fluctuations
- We would like to propagate this uncertainty into the derived TPR
  - Binomial errors on n_i, p_i
  - Derive uncertainty on TPR from MC
- In addition to the best-fit TPR, set a 95% Confidence Level (C.L.) lower limit on the TPR



Galaxy Clusters

- SVA1 gold catalog objects within 0.5 deg of 11 redmapper clusters with >100 members
- Separate objects based on the modest star-galaxy classifier in SV-A1
- Define regions:
  - region 1: 0 < alpha < 0.005 deg²
  - region 2: 0.20 < alpha < 0.205 deg²
- Angular distributions look well behaved, so try to apply the formalism developed previously
  - Best-Fit TPR (completeness): 94.5%
  - Lower Limit on TPR (95% C.L.): 92.8%





Conclusions and Caveats

- **This method only incorporates statistical uncertainties.**
- **De-blending is an issue!**
  - **If the we aren't finding stars near the center of galaxy clusters, this will lead to an over-estimation of completeness**
  - **Dense stellar clusters are currently heavily impacted**
- **Cluster selection:**
  - **It would be nice to have a selection of clusters that don't rely on the analysis of the DES data (i.e., and ad hoc list rather than an a posteriori list from redmapper).**
- **I'm in the process of writing a note to detail the formalism.**