

CONTINUING ANALYSIS OF WHITTIER COLLEGE ADMISSIONS DATA: EXAMINING MATRICULATION TO SOPHOMORE YEAR

Jordan Hanson

December 6th, 2018

Whittier College Department of Physics and Astronomy

OUTLINE

1. **Motivation:** *Continuation of the study of admissions data outliers and the aid gap distribution*
2. **Methods:** *Derive normalized variables from raw data*
3. **Results:**
 - *There are outliers who do not proceed to sophomore year.*
 - *The aid gap distribution is bi-modal, and one of the two sub-distributions contains the students that do not proceed.*
4. **Next steps:** *Correlation of aid gap data with proceeding to sophomore year*
5. **Conclusion**

MOTIVATION

MOTIVATION: RESULTS FROM PREVIOUS PRESENTATION

- Students with very low scores are being admitted.
- Attempted to classify students into two groups:
 - Those who will proceed to sophomore year
 - Those who will not proceed to sophomore year
- Further findings:
 - Admissions variables are correlated for students that proceed to sophomore year
 - Admissions variables are uncorrelated for students that do not proceed to sophomore year
 - **Example:** A student who has one high SAT score but two low ones will probably not graduate.
 - **Example:** A student who has a high GPA in high school but a low GPA in the first semester at Whittier is at risk of not graduating.

METHODS

1. **Selection:** *From raw data:*

- Require three SAT scores, and high school GPA
- Require GPA from Whittier, freshman fall and spring
- Require INTD 100 grade
- Require aid gap knowledge

2. **Methods:** *Derive normalized variables from raw data*

Normalizing data: Let s_i represent a score (like GPA) for student i , and let \bar{s} and σ_s represent the mean score and standard deviation for the students, respectively. The normalized score is

$$s_{i,n} = \left(\frac{s_i - \bar{s}}{\sigma_s} \right) \quad (1)$$

Criteria	N
Raw data	3119
Require parameters	1346

Table 1: The numbers of students in the analysis. The student data corresponds to 2012-18.

Parameter	Mean	Std. dev.
SAT CR	525.6	82.1
SAT Math	526.5	82.5
SAT Writing	523.0	77.7
GPA HS	3.52	0.42
GPA FF	3.00	0.72
GPA FS	3.01	0.67
INTD100	3.20	0.89
Aid Gap	539.2	14560

Table 2: The numbers of students in the analysis. The student data corresponds to 2012-18.

RESULTS: GPA OUTLIERS

RESULTS: OUTLIERS DO NOT PROCEED TO SOPHOMORE YEAR

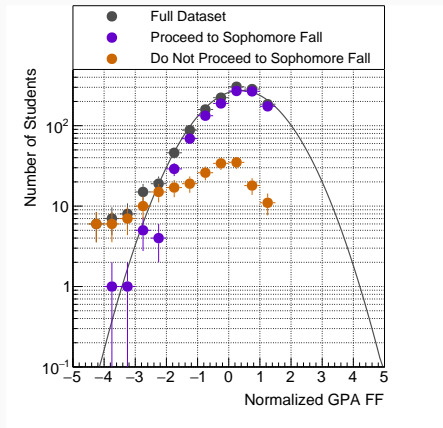


Figure 1: The normalized GPA for the first fall of Whittier College students. The mean GPA in the first fall is 3.00, with a standard deviation of 0.72.

RESULTS: OUTLIERS DO NOT PROCEED TO SOPHOMORE YEAR

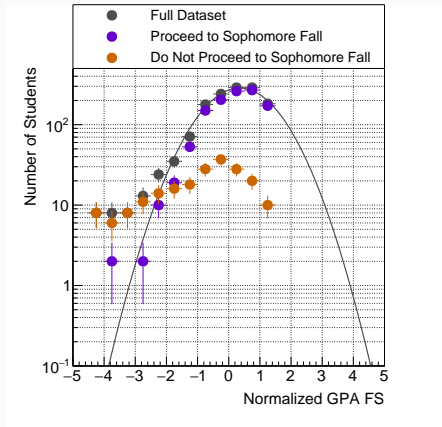


Figure 2: The normalized GPA for the first spring of Whittier College students. The mean GPA in the first spring is 3.01, with a standard deviation of 0.67.

RESULTS: AID GAP DISTRIBUTIONS

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

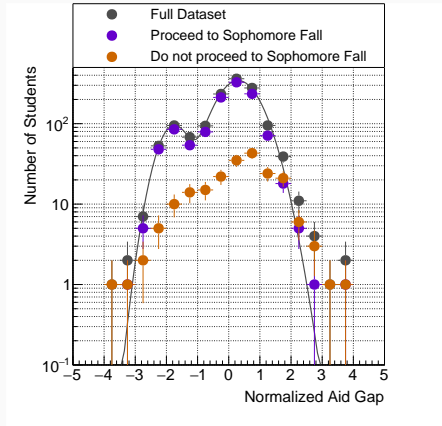


Figure 3: A coarsely binned aid gap distribution is fit with a double-Gaussian distribution.

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

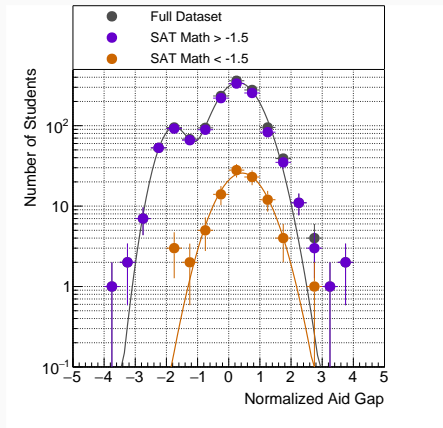


Figure 4: The positive aid gap data corresponds to Math SAT scores below about 1.5 (mean=526.5, std=82.5).

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

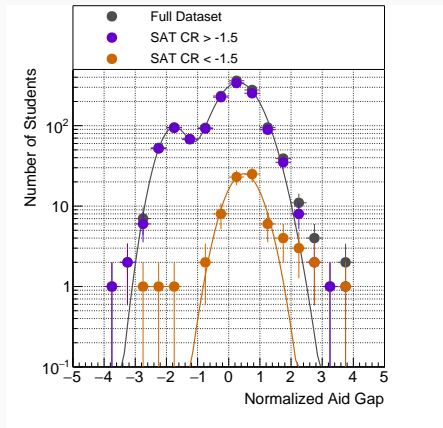


Figure 5: The positive aid gap data corresponds to CR SAT scores below about 1.5 (mean=525.6, std=82.1).

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

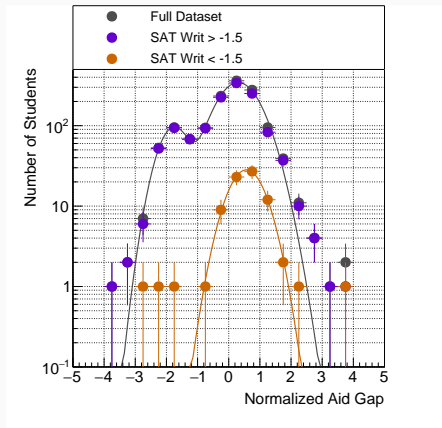


Figure 6: The positive aid gap data corresponds to Writ SAT scores below about 1.5 (mean=523.0, std=77.7).

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

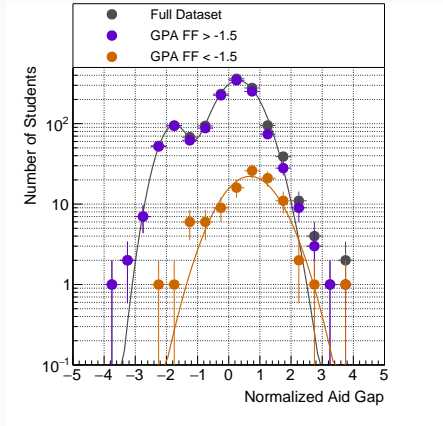


Figure 7: The positive aid gap data corresponds to GPA FF scores below about 1.5 (mean=3.00, std=0.72).

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

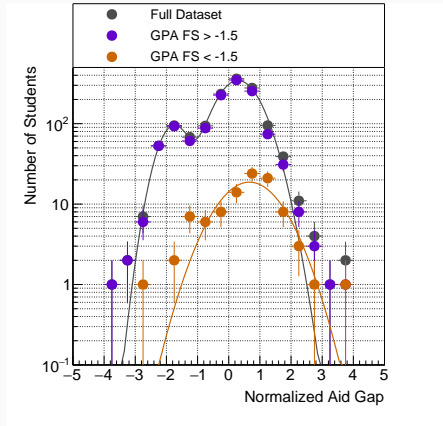


Figure 8: The positive aid gap data corresponds to GPA FS scores below about 1.5 (mean=3.01, std=0.67).

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

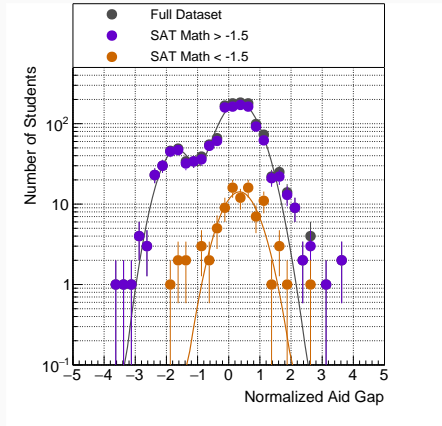


Figure 9: The two-distribution effect still appears after finer binning.

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

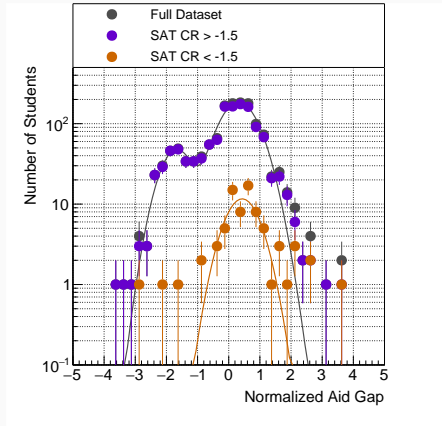


Figure 10: The two-distribution effect still appears after finer binning.

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

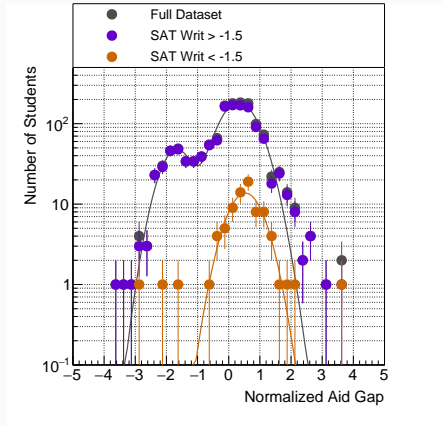


Figure 11: The two-distribution effect still appears after finer binning.

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

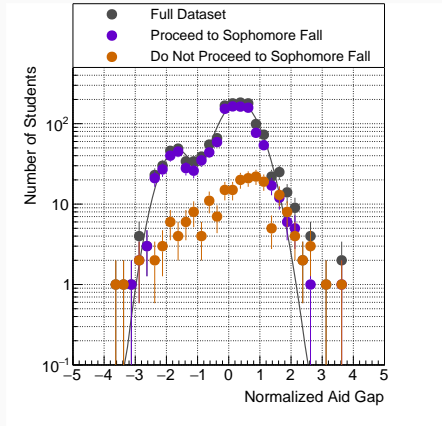


Figure 12: The two-distribution effect still appears after finer binning.

RESULTS: AID GAP DISTRIBUTION IS REALLY TWO DISTRIBUTIONS

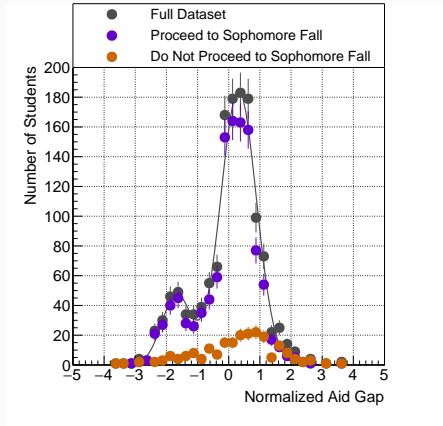


Figure 13: The aid-gap distribution without the logarithmic scale.

RESULTS: CUMULATIVE DISTRIBUTION FUNCTIONS (GPA FF)

RESULTS: CUMULATIVE DISTRIBUTION FUNCTIONS (GPA FF)

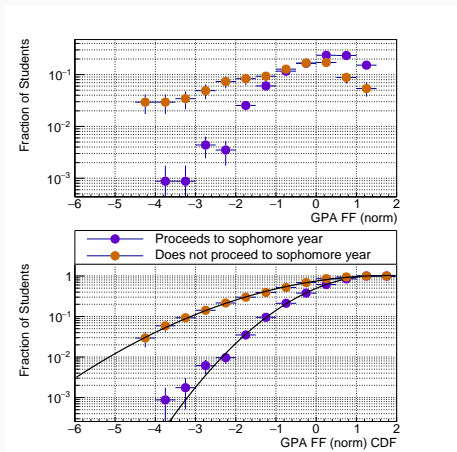


Figure 14: The PDFs (top) and CDFs (bottom) for students who proceed (purple points) and do not proceed (orange points).

CONCLUSIONS

CONCLUSION

1. **GPA outliers:** *Students with very low GPA occur in our population at a rate much larger than expected from standard statistics.*
2. **GPA outliers:** *These students will not proceed to sophomore year.*
3. **Aid Gap distributions:** *The aid gap distribution is comprised of two sub-distributions.*
4. **Aid Gap distributions:** *If a student has low SAT scores or low GPA, that student is more likely to belong to the positive aid gap sub-distribution.*
5. *Let's be careful about inferences, because not every variable is correlated. For example, HS GPA is messy.*