Does visiting campus affect prospective students' decision to attend TCU?

Executive summary:

Yes. Very much yes. If there weren't a relationship between visiting and depositing, and it was a random effect causing visitors to deposit more often, this random effect would be more likely than every person on the planet winning the lottery at the same time every day for the next 14 years. Visiting campus obviously isn't the only thing that influences an applicant to make their deposit, but those who do visit are statistically much more likely to do so. So let's keep up the on-campus programs and give students the best experience we can when they are here.

Background:

I am looking to see if there is a relationship between visiting campus and depositing. Many of us probably feel that there is a relationship! It seems obvious that potential students (admitted or prospective) who take the effort to visit campus are more likely to enroll than those who do not visit. One of my biggest passions is putting evidence and statistics behind our gut feelings.

For this study, I am using the **chi-squared test of association** to determine if there is a statistically significant relationship between two categorical variables. A *categorical variable* is a variable that may take on one of a set of labels. For example, gender: male/female/other; visited campus: yes/no; vehicle owned: Toyota/Honda/etc.. The Chi-Squared test is a statistical hypothesis test which assumes that the *observed* occurrences for a categorical variable match the *expected* occurrences for the categorical variable, if they don't match, then there could be a relationship between the two variables. The expected occurrences are computed based on prior experiments, institutional knowledge, or other outside methods.

Data and methodology:

The data I will use are all of the most-recent freshman applicants from 2016-2018. That is, for any applicant with more than one application for any of 2016, 2017, or 2018, I only consider their "rank 1" application which is their most recent one. Furthermore, I excluded 175 records who visited campus *after* depositing. This leaves us with about 59,000 records to analyze.

I will actually present two results. One testing the overall applicant pool described above, and another just restricting the results to admitted students.

The following table shows the counts of each student from each category:

OVERALL APPLICANTS - Observed	NO Deposit	YES Deposit	Totals
NO Visit	40204	1531	41735
YES Visit	12790	4507	17297
Totals	52994	6038	59032

Figure 1 – Contingency table for overall applicant pool

The following table shows the counts of each student from each category, but it's limited to only applicants who have been admitted:

ADMIT APPLICANTS - Observed	NO Deposit	YES Deposit	Totals
NO Visit	10313	1531	11844
YES Visit	7133	4507	11640
Totals	17446	6038	23484

Figure 2 – Contingency table for admitted students only

In 2016, there were 1888 first year freshman deposits from 7506 admits of 19972 apps In 2017, 1955 deposits from 8110 admits of 19740 apps, and finally in 2018, 2194 deposits from 8210 admits of 20156 apps. Therefore the expected percentage of depositors should be:

$$Expected\ yield\_admits = \frac{1888 + 1955 + 2194}{7506 + 8110 + 8210} = \frac{6037}{23826} = 0.25337866$$

Or, about 25.3% of admitted students should deposit, and:

$$Expected\ yield\_prosp = \frac{1888 + 1955 + 2194}{19972 + 19740 + 20156} = \frac{6037}{59868} = 0.10083841$$

About 10.1% of all applicants should deposit. From this we can get the expected values for each contingency table. The expected percentage is multiplied by the number of people who visited we expected to deposit, and one minus the expected percentage is used to get the number of people we expect not to deposit.

The next step is to compare the actual counts to the expected counts, then perform a series of calculations to arrive at a value, called chi (the Greek letter χ).

Results:

OVERALL APPLICANTS - Expected	NO Deposit	YES Deposit	Totals
NO Visit	37526	4208	41734
YES Visit	15552	1744	17326
Totals	53078	5952	59030

Figure 3 – Expected Contingency table for overall students

ADMIT APPLICANTS - Expected	NO Deposit	YES Deposit	Totals
NO Visit	8842	3001	11843
YES Visit	8690	2949	11639
Totals	17532	5950	23482

Figure 4 – Expected Contingency table for admitted students only

The above tables show the expected (theoretical) counts of individuals belonging to those categories. The totals may not match exactly due to rounding of decimals. As an example, for the 23482 admitted students, we'd expect 5950 (about 6k) to deposit. Since admitted students are almost equally likely to visit or not, we would expect about 3k to deposit without visiting, and 3k to deposit after visiting.

Generally speaking, if there is less than 5% chance for a completely random effect to explain the difference between the expected value and the actual value, it would be considered statistically significant; especially when the experiment involves humans. Some experiments in physics involving trillions of data points require less than 0.000001%, five zeros after the decimal, attribution to random effects to be considered significant.

In this study, if the relationship between visiting campus and depositing were to be explained by a random effect, it would have a probability of 0.000000000… with AT LEAST 293 zeroes later! I say *at least* because my computer literally can't display a number that small.

Conclusion:

This is a mind-bogglingly significant difference! Consider this, we expected 3k admitted students to deposit without visiting, and 3k to deposit after visiting. However, we got 4.5k from the visitors and 1.5k from the non-visitors instead! If this difference was because of a random effect, TCU would have to force every one of our future admitted students to visit our campus every year for about 50 times longer than the universe has existed before seeing this result again.

In summary, visiting campus obviously isn't the only thing that influences an applicant to make their deposit, but those who do visit are statistically much more likely to do so. So let's keep up the on-campus programs and give students the best experience we can when they are here.

Thank you,

Michael Greene, PhD