Day 9

Outline

1. Conditional probability example
2. Power analysis example

Please reference the attached sheets for full examples

Conditional Probability Examples

Example 1

In a lecture class of 150 students, 110 students are freshmen, 50 own a dog, and 25 are freshmen who own a dog. Suppose a student is selected at random.

Tree Diagram Version

Root

- Freshman ($\frac{110}{150}$)
  - Own dog : $\frac{25}{110}$ [Freshman AND own dog = ($\frac{110}{150} \times \frac{25}{110} = \frac{1}{6}$)]
  - No dog : $\frac{85}{110}$ [Freshman AND no dog = ($\frac{110}{150} \times \frac{85}{110} = \frac{17}{30}$)]
- Not Freshman ($\frac{40}{150}$)
  - Own dog : $\frac{25}{40}$ [Not freshman AND own a dog = ($\frac{40}{150} \times \frac{25}{40} = \frac{1}{6}$)]
  - No dog : $\frac{15}{40}$ [Not freshman AND not own a dog = ($\frac{40}{150} \times \frac{15}{40} = \frac{1}{10}$)]

a. What is the probability of being a freshman, given that the student owns a dog?

- $P(Freshman|Dog) = \frac{P(Freshman\, AND\, Dog)}{P(Dog)} = \frac{\frac{25}{150}}{\frac{25}{150} + \frac{25}{150}} = \frac{25}{50} = \frac{1}{2}$

b. What is the probability of owning a dog, given that the student is a freshman?

- $P(Dog|Freshman) = \frac{P(Dog\, AND\, Freshman)}{P(Freshman)} = \frac{\frac{25}{100}}{\frac{110}{150}} = \frac{25}{110} = \frac{5}{22}$

- P(Freshman & Dog) = P(Freshman) P(Dog|Freshman)
- P(Freshman & Dog) = P(Freshman) P(Dog) $\leftarrow$ Independence

Table Diagram Version

	Freshman	Not Fresh.	Total
Dog	25	25	50
No Dog	85	15	100
Total	110	40	150

Figure 1: Freshman Table Example

Power Analysis Examples

Example 1

It is believed that about 10% of the population is left-handed. However, China has claimed that less than one percent of its students are left-handed. Suppose we are interested in evaluating whether there is something special about Chinese people, or whether the Chinese government is lying. Suppose further that we have devised a scientifically perfect test to measure a person's dominant hand. Would a random sample of 50 Chinese students be large enough to detect a population difference of 10% vs. 1%?

We want low α and high power

- $\mathrm{H_0}$: p = 0.1 [Null]
- $\mathrm{H_1}$: p = 0.01 [Alternate]
- `N` : 50
- α: 0.05
  - If α is not given, please assume $\alpha = 0.05$
- Define p = proportion of left handed students
- For midterm one, define X = number of (successes) left-handed students in our sample.
- Decision rule:
  - Critical region: $X \leq x$
  - If X is in critical region, accept $\mathrm{H_0}$, else accept $\mathrm{H_1}$.
  - Only problem is we <u>do not</u> know what `x` is.
  - Defining our critical region to be $X \leq 5$
    - Under the null hypothesis $\mathrm{H_0}$, $X \sim B(50, 0.1)$
      - $\mathrm{P}(X \leq 5 | p = 0.1) = 0.616$
      - P(Type 1 Error) = 0.616
    - Under the alternative hypothesis $\mathrm{H_1}$, $X \sim B(50, 0.01)$
      - $\beta = P(X > 5 | p = 0.01) = 0$
      - Power $= P(X \leq 5 | p = 0.01) = 1$
  - Defining our critical region to be $X \leq 1$
    - When p = 0.1
      - 3.4% false positive
      - 96.6% true negative
      - $\alpha = 0.034$
    - When p = 0.01
      - 91.1% true positive
      - 8.9% false negative
      - Power = 0.911
      - $\beta = 0.089$

Example 2

Is this sample large enough to detect something → power rule!!!!!!!

We want low α and high power

- H_0 : p = 0.26 [Null]
- H_1 : p = 0.52 [Alternate]
- `N` : 14
- α: 0.05
  - If α is not given, please assume $\alpha = 0.05$
- Define p = % of patients progression free after 6 months

Using `R`:

- Critical region is X > 6 or $X \geq 6$
- `lower.tail = TRUE` includes $\leq$
- `lower.tail = FALSE` includes >
- When P = 0.26
  - 4.7% false positive
  - 95.3% true negative
  - $\alpha = 0.047$
- When P = 0.52
  - 66.2% true positive
  - 33.8% false negative
  - Power = 0.662
  - $\beta = 0.338$