Evaluation of Mathematical Models

In what ways can a model be "good"? A model can be...

- **Accurate**
  - Is the output of the model very near to correct?
- **Descriptively Realistic**
  - Is the model based on assumptions which are correct?
- **Precise**
  - Are the predictors of the model definite numbers?
- **Robust**
  - Is the model relatively immune to errors in the input data?
- **General**
  - Does the model apply to a wide variety of situations?
- **Fruitful**
  - Are the conclusions useful?
  - Does the model inspire other good models?

Accuracy

Accurate
Real
Precise
Robust
General
Fruitful

Definition: A model is **accurate** if the answers it gives are correct.

Example. Determining projected student populations.

This year, there are 10 million people between 18–22 years old. (P)

This year, there are 5 million students this year. (S)

We might conjecture that in general, $S = 0.5P$.

Model Assumption 1:

Model Assumption 2:

If next year there are projected to be $11,000,000$ 18–22 year olds, we would estimate the college population to be of size $E =$ ______.

If this value is close to correct, we say our model is accurate. Otherwise, the model is **inaccurate**.

Problem: (We won't know this until next year!)

Question: Is this model descriptively realistic?

Descriptively Realistic

Accurate
Real
Precise
Robust
General
Fruitful

Definition: A mathematical model is **descriptively realistic** if it is deduced from a believable description of the system being modeled.

Example. **Full moons.** A full moon appears to occur every 29 days. Let M_L, M_N be the dates of the last and next full moons. Is the model

$$M_N = M_L + 29$$

descriptively realistic? _____ Why?

Descriptively Realistic

Accurate
Real
Precise
Robust
General
Fruitful

Example. A more descriptively realistic model would incorporate other age groups. Replace Assumptions 1 and 2 by:

Model Assumption 3: College students are either:

- 18–22 (P_a of these)
- 23 or older (P_b of these)
- 17 or younger (P_c of these)

Model Assumption 4: The enrolled percentages for each age range is:

- 30% for people aged 18–22
- 3% for people aged 23 or older
- 1% for people aged 17 or younger

We would estimate the college population to be of size

$$E = 0.3P_a + 0.03P_b + 0.01P_b.$$

Precision

A model is
- **precise** if the prediction is
  - a definite number
  - a definite function
  - etc.
- **imprecise** if the prediction is
  - a range of numbers
  - a set of functions
  - etc.

Circle one: The above models are $\boxed{\text{precise} \quad \text{imprecise}}$. Why?

Keep *Assumption 1:* Each college student is in 18–22 year old range.

Revise *Assumption 2*:* The percentage of 18–22 year olds in college is between 46% and 50%. (Historically)

Model Conclusion:
$$(0.46)(11,000,000) \le E \le (0.5)(11,000,000)$$
$$5,060,000 \le E \le 5,500,000.$$

This model is imprecise, but perhaps more helpful than the precise answer from before.

Robustness and Percentage Error

Accurate
Real
Precise
Robust
General
Fruitful

Definition: A model is **robust** if it is relatively immune to errors in the input data.

$$\boxed{\text{Input Error}} \xrightarrow{\textit{model}} \boxed{\text{Output Error}}$$

Example. If our population estimate (input) has an error of 10%, how much does our college enrollment estimate (output) change?

Ask: Is the output error less than 10% or more than 10%?

- Some models **magnify** the errors that exist in the input data; we say these models are **sensitive to error** or **not robust**.

Make sure we understand: What does 10% error mean?

Percentage Error

Accurate
Real
Precise
Robust
General
Fruitful

Definition: Suppose you are finding the value of something. Let v be its true value and v' be the value predicted by a model or measured.

- The **error** is calculated by $v' - v$.
- The **fractional error** is calculated by $\dfrac{v' - v}{v}$.
- The **percentage error** is calculated by $\left(\dfrac{v' - v}{v} \cdot 100\right)\%$.

Example. Suppose that the census measures the 18-22 year old population to be $9,300,000$ while the true population is $9,500,000$.

- The **error** is
- The **fractional error** is
- The **percentage error** is

Most of the time, we discuss the *absolute value* of percentage error.

In other words, 5% error means the error is either -5% or 5%.

Percentage Error

Example. How robust is our $E = 0.5P$ model?

Suppose that we prepare for a $+5\%$ error in population.

Recall: Population Estimate $P' = 11,000,000$.
Calculating the true population P based on a $+5\%$ error in P':

$$\frac{11{,}000{,}000 - P}{P} = 0.05 \implies 11,000,000 - P = 0.05P \implies$$
$$11,000,000 = 1.05P \implies P = 10,475,190$$

Note: The true population P is **less** than the estimate P' because our estimate was 5% **too high**.

How does this impact the true student enrollment E?

$$E = 0.5P = 0.5(10,475,190) = 5,238,095,$$

which is an error of: $\frac{5{,}500{,}000 - 5{,}238{,}095}{5{,}238{,}095} = 0.05$

This highlights the principle of "Error In equals Error Out"

Percentage Error

Accurate
Real
Precise
Robust
General
Fruitful

Example. How robust is our $E = 0.3P_a + 0.03P_b + 0.01P_c$ model?

Suppose that we prepare for a $\pm 10\%$ error in each population P_i, where the true values are: $P_a = 10$ mil., $P_b = 90$ mil, $P_c = 50$ mil.

If each pop. est. P_i is a 10% **overestimate** of the true value P_i', $P_a' = 11$, $P_b' = 99$, and $P_c' = 55$.

Then comparing the true enrollment to the estimated enrollment E':

$$\begin{aligned} E \; &= 0.3(10) + 0.03(90) + 0.01(50) = 6.2 \\ E' &= 0.3(11) + 0.03(99) + 0.01(55) = 6.82 \end{aligned}$$

Percentage error: $\frac{6.82-6.2}{6.2} = \frac{.62}{6.2} = 10\%$; Again ____________

Alternatively, P_a' 10% **underestimate**, and P_b', P_c' 10% **overestimate**: $P_a' = 9$, $P_b' = 99$, and $P_c' = 55$.

$$\begin{aligned} E \; &= 0.3(10) + 0.03(90) + 0.01(50) = 6.2 \\ E' &= 0.3(\;\,9) + 0.03(99) + 0.01(55) = 6.22 \end{aligned}$$

Percentage error: $\frac{6.22-6.2}{6.2} = \frac{.02}{6.2} = 0.3\%$.

Generality

Accurate
Real
Precise
Robust
General
Fruitful

Definition: A model is **general** if it applies to a variety of situations.

Model Assumption: Each college student is in 18–22 year old range.

Model Assumption: Each college will have its enrollment change by the same ratio, next year's 18–22 year old population over this year's.

Suppose that Queens College has 20,000 students and suppose that Private UNnamed Kansas College has 2,000 students this year.

If the year-to-year change in 18–22 year old population is 10%, then QC would gain 2,000 students while PUNK College would gain 200.

The projected enrollment in all colleges would be:

$$\begin{aligned} E &= (1.1)S_1 + (1.1)S_2 + \cdots + (1.1)S_n \\ &= (1.1)(S_1 + S_2 + \cdots + S_n) \\ &= (1.1)S \end{aligned}$$

It is complicated to estimate total enrollment using this model.
This model is more general because it applies to individual colleges.

Fruitfulness

Accurate
Real
Precise
Robust
General
Fruitful

Definition: A model is **fruitful** if either

- Its conclusions are useful.
- It inspires other good models.

Our college enrollment model is fruitful in multiple ways:

- Planning for demand for educational grants, dormitory space, teacher hiring, etc.
- The ideas we implemented are transferrable to other situations.

Example. How many automobiles would be junked in a given year?

- Cars play the role of people.
- Partitioning by age of cars gives better results

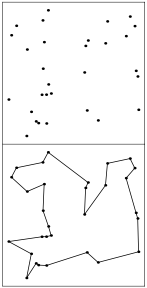
The Advantage of Inaccuracy

Accurate
Real
Precise
Robust
General
Fruitful

Often accuracy is very expensive (either computationally or financially).

Example. The Traveling Salesman Problem

TSP: Given a home location and a set of places to visit, find the shortest path that starts and ends at home and visits each of the places along the way.



With many locations, there are (inexpensive and inaccurate) or (expensive and accurate) algorithms to solve these problems.

Your approach will depend on the particular application and your scale:

- If you visit the same places every day, run the expensive model **once initially** in order to save money in the long run.
- If you visit different places every day, run the inexpensive algorithm daily. (Unless you're UPS or FedEx.)