1/28 Math 390.4 lecture 1

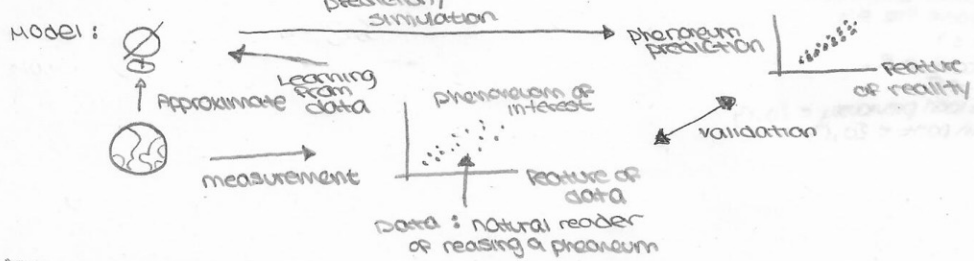
"Models" are approximations
or "approximations" to
"reality", "truth", "phenomeum

Model	≈	Reality
Model airplane		Airplane
Map		Roads in a city
Windtunnel		Fast moving air in the sky
"Early to bed, early to rise, makes a man wealthy"		success of a person

"All models are wrong, but some are useful"
- George Box and Draper
(approximate, purposeful)

⊕ radius: 1
$C \neq 3.14 \Rightarrow C \approx 3.14$

Two main goals:
1) can the model predict a future event of the phenomeum under examination (prediction)
2) How does the phenomeum work? (explanation) (goal of science)



"Early to bed, early to rise, makes a man healthy, wealthy, and wise."
} analogous

What is this modeling?
1) Health
2) Wealth
3) Wisdom
} 3 different phenomeum

To make the model concrete, we need numerical definitions
We need "metrics", those define how to measure both phenomeum and features of reality.

What are the features?
1) Bedtime
2) wake-time

Metric for:
Bedtime:
Avg. Bedtime in 24 hr time (Yes, Yes, Yes, no)
Avg. Bedtime in seconds and starts from 17:00 from age 18 - 65

Metric evaluation
1) does it capture the feature / phenomeum?
2) is it easily reproducible and unambiguous?
3) Good resolution?
4) Monitanic?

Bedtime (b)
wake-time (w)
health measured by longevity (l)
wealth measured by net worth at 65 (m)
wisdom measured by test (s)

we want to estimate F where:

$$\left[\frac{h}{s^2}\right] = F(b, w)$$

prediction of phenomeum ← mathematical model (ideas and abstractions, not physical entities)

examples:
$a = \frac{F}{m}$ $F(m, F)$; $E = mc^2$

$Y = t(Z_1, Z_2, \ldots, Z_t)$

The phenomenon of response, outcome, signal, endpoint or dependent var. (1 dimensional)

True causal inputs

The function (unknown) that combines $Z_1 \ldots Z_t$

Phenomenon

Pay back mortgage ($y = 1$)
or not pay back mortgage ($y = 0$)

$y \in \{0,1\}$ = Y (output space) (binary)
↑ positive class

What are the causal inputs?

$Z_1, Z_2 \ldots Z_t$

Z_1: has the money at payback time $\{0,1\}$

Z_2: unforseen emergency $\{0,1\}$

Z_3: criminal intent $\{0,1\}$

$t(Z_1, Z_2, Z_3) = Z_1(1 - Z_2)(1 - Z_3)$

Modeling Problem

When you don't know the Z's of t

Next best thing:

Obtain measurements that approximate the Z's called (X's)

X_1: credit score $\in \mathbb{R}_+$

X_2: salary $\in \mathbb{R}$

X_3: missed loan previously $\in \{0,1\}$

X_3: crime in past $\in \{0,1\}$