Credit Worthiness Model

$y \in \{0,1\} = y$

we had a true system (not a model but sometimes called the "true model")

$y = t(z_1, z_2, z_3)$

z_1: has sufficient funds
z_2: unforeseen emergency
z_3: criminal intent

Big Problem: $\{z_1, z_2, z_3\}$ are unobservable

Impossible: get $\{z_1, z_2, z_3\}$ t

Next Best: try to define & collect information "related" to the true causal inputs

x_1: Salary ; How To measure it?
↳ Take Average over 5 yrs. $\in \mathbb{R}$

x_2: previous loan repayment $\in [0,1]$

x_3: previous crime type $\in$ {no crime, infraction, felony}
↳ issue: not a number(s)

Bob's Information: $\vec{x} = [x_1, x_2, x_3] \in \mathcal{X}$

$\vec{x}$: observation/record/object/input/independent variables

$[x_1, x_2, x_3]$: features AKA variables, characteristics, attributes, covariances, predictors in Economics: regressor

$\mathcal{X}$: The Covariance Space

$\dim[\vec{x}] = p$ or "d" where p is how many features $\vec{x}$ has

$x_1 \in \mathbb{R}$: "continuous variable"

$x_2 \in \{0,1\}$: "binomial or dummy variable"

x_3 is a "categorical variable" w/ 4 "levels" ↳ unique possible values

In the case of x_3 we can represent it as:

$$x_3 \in \{0,1,2,3\}$$

Though this should only be done if the categorical predicator is "ordinal".

A solution is to convert x_3 into binary features:

$x_3 \rightarrow x_{3A}, x_{3b}, x_{3C}, x_{3d}$

- x_{3A}: binary no crime
- x_{3b}: binary infraction
- x_{3C}: binary misdemeanor
- x_{3d}: binary felony

So lets say x_{3C} = misdemeanor, our $\vec{x} = [\ldots 0010]$

Notice $\dim[\vec{x}] = 6$ with the addition of $x_{3A}, \ldots, x_{3D}$

i.e. there is an increament in the size of the dimension

Impossible: get $\{z_1, z_2, z_3\}$ but we have $[x_1, x_2, x_3]$

Goal: do the best we can in explaining y by creating model f, the approximation is the best functional relationship we can get

Does $y = f(x_1, x_2, x_3)$ ever? ⟶ NO

Instead

$y \approx f(x_1, x_2, x_3)$

$y = f(x_1, x_2, x_3) + \delta$ ← comes from ignorance

where $\delta = t(\vec{z}) - f(\vec{x})$

How do we get f ?

First note there is <u>no analytical solution</u>.

e.g.: $h(x) = x^2$ find $\min\{h\}$

Instead use an "empirical solution" i.e. use data

→ "Learning from data"

- "Supervised Learning" use historical examples of records & their responses.
  - It requires 3 ingredients:
    1. $\mathcal{D} := \{\langle \vec{x}_1, y_1 \rangle, \langle \vec{x}_2, y_2 \rangle, \ldots, \langle \vec{x}_n, y_n \rangle\}$

$\vec{x}_1$ is Bill's characteristics & y_1 is whether or not he paid back loan

$\vec{x}_2$ is Jill's " " " " " & y_2 " "

Thus Let $X = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_n \end{bmatrix} \in \mathcal{X}^n$

$\mathcal{D} = \langle X, \vec{y} \rangle$

$\dim(X) = n \cdot p$

$\dim(\vec{y}) = n$

$\vec{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathcal{Y}^n$