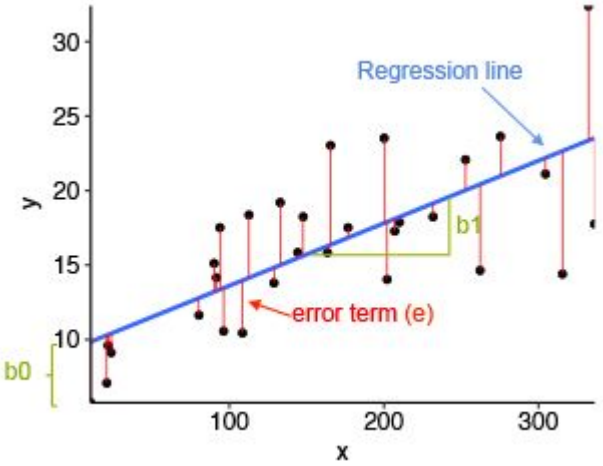
TUT206 OCT25

Announcement!

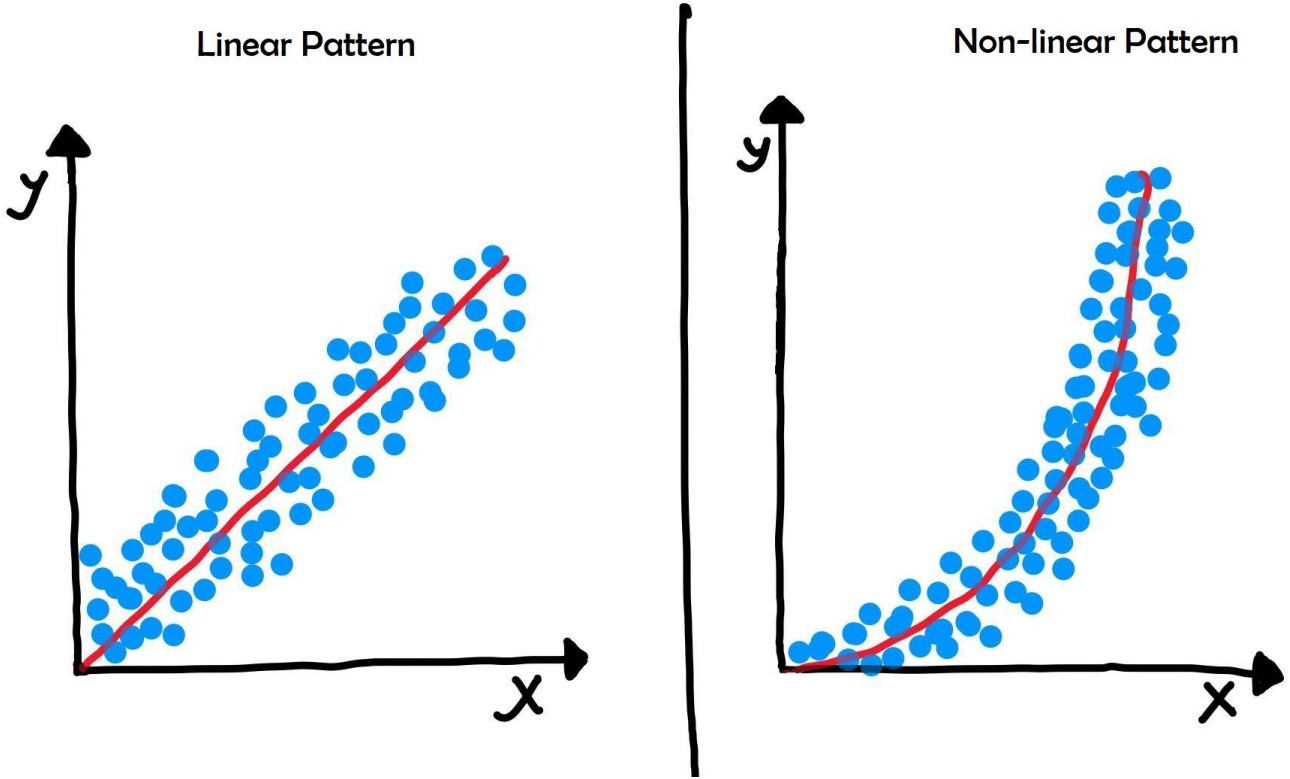
- The course project link
- **The MIDWAY ChatBot Experience FEEDBACK** link
- Interaction? See github md

Recap: simple linear regression



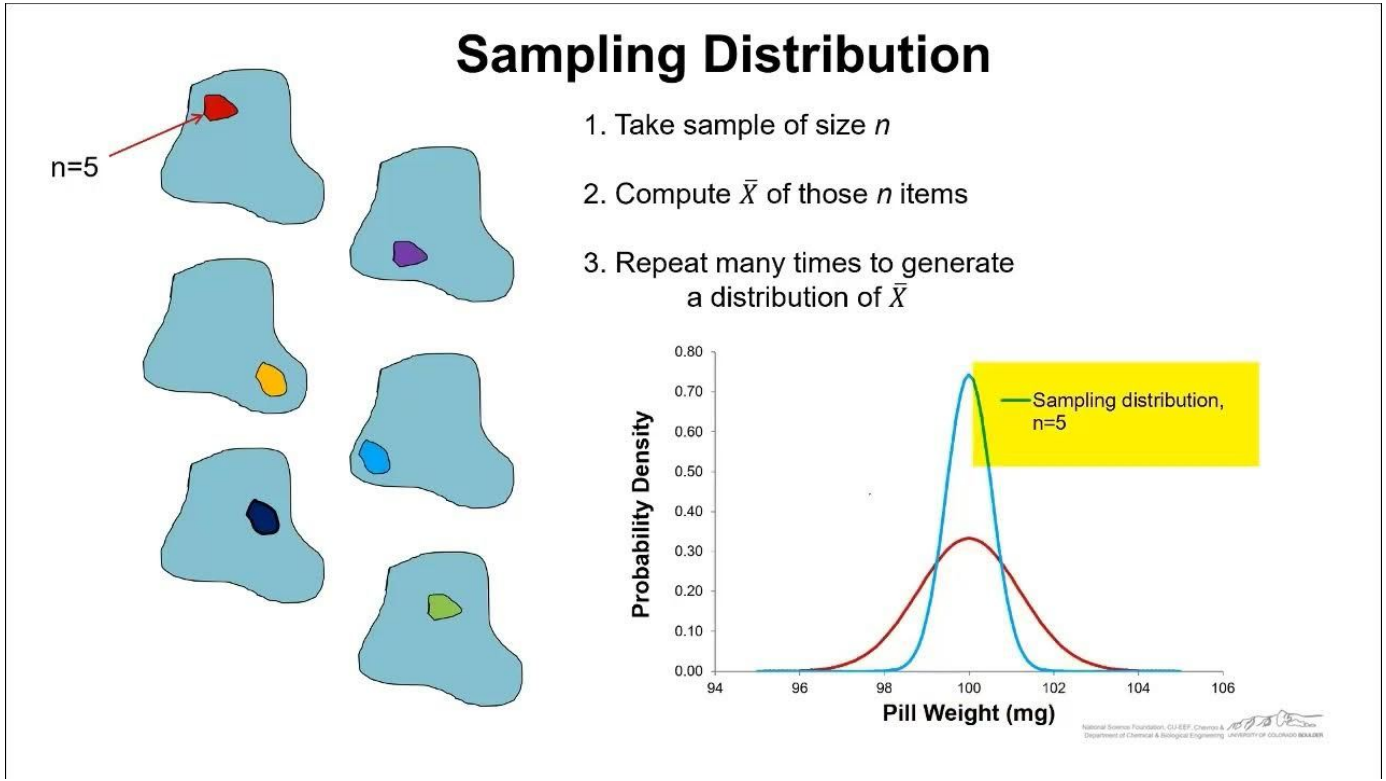
$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \text{where} \quad \epsilon_i \sim \mathcal{N}\left(0, \sigma^2\right)$$

Recap: simple linear regression assumption



Recap: sample distribution! important

[Sampling Distribution vs Population Distribution (youtube.com)]



Recap: sample distribution! important

HW this time is Going To Be DIFFERENT: you MUST understand simulation to do it

AKA **Hypothesis Testing**, **Sampling Distribution under the Null Hypothesis**, and related topics regarding interpretation from Oct04 TUT and Oct11 TUT; AND, **Sampling Distribution**, **Bootstrapped Confidence Intervals**, and related topics regarding interpretation from Sep27 TUT and Sep30 LEC and Oct07 LEC

Show how to using notebooklm

Communication:

The Wheel of Destiny

Stella McStat had been running a small-time gambling operation on campus for several months during her first year at UofT...

- For each spin of the wheel, two gamblers take part. For a toonie each (\$2 Canadian), Stella sells one a red ticket and one a black ticket (i.e., total \$4). Then Stella spins the Wheel of Destiny. The person who holds the colour on which the spinner stops gets \$3.50 (Stella keeps \$0.50 per spin for running the game and providing snacks).

Communication

Communication

1. **Null Hypothesis (H_0)**:
    - In the context of "Stella McStat's Wheel of Destiny," the **Null Hypothesis** (H_0) could be that the **wheel is fair**. This means that the outcomes of the wheel occur at equal probability (e.g., no section of the wheel is favored over another).
2. **Alternative Hypothesis (H_1)**:
    - The **Alternative Hypothesis** (H_1) is that the **wheel is biased**. In this case, some outcomes may be more likely to occur than others, indicating that the wheel is not fair.
3. **P-Value Definition**:
    - A **p-value** is the probability of obtaining a test statistic as extreme as (or more extreme than) the observed test statistic, assuming that the **null hypothesis** is true.

Communication:

Here's a possible example of x and y values:

- x: The size of each section on the wheel, measured in degrees or as a percentage of the wheel (e.g., 30%, 20%, 10%, etc.).
- y: The number of times the wheel lands on each section during a series of spins.

Communication:

Communication

In the context of **simple linear regression**, the **Null Hypothesis (H_0)** typically states that the **slope** $\beta_1 = 0$. This implies that there is no relationship between the predictor variable x and the response variable Y, meaning that changes in x do not affect Y.

The **Alternative Hypothesis (H_1)** is that $\beta_1 \neq 0$, which implies that there **is a relationship** between x and Y, meaning that changes in x are associated with changes in Y.

Mathematically, these hypotheses are:

- **Null Hypothesis (H_0)**: $\beta_1 = 0$
- **Alternative Hypothesis (H_1)**: $\beta_1 \neq 0$

Demo: concepts see zoom

Demo: Observed data setup

Show chatgpt canvas

Demo: Model fitting code demo

```
statsmodels.formula.api as smf
```

```
smf.ols(...).fit()
```