Simulation & Resampling

For MC Sampling → you assume a theoretical model

Ex 1

- Let x = # of times indices ↑ in a year
- mean(x) (+ possibly, var(x)?)

a) $\bar{x} = 75\%$ (pointwise estimate)

b) $[\bar{x} - v, x + v]$ (e.g.) $(75\% - 10\%, 75\% + 10\%)$

↳ If you know your variability in the model then you can manage your risk

Ex 2

# of babies born per week

- $\bar{X}$ is a R.V since it depends on x (which is a R.V)
- $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ (CLT)

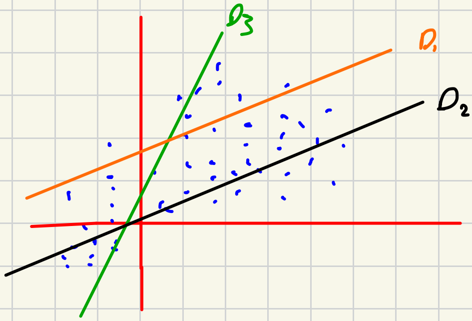
Assuming x_i iid with $E[x] = \mu$ and $\text{var}(x) = \frac{\sigma^2}{n}$

Ex

You want to use $Y = \beta x + \varepsilon$ on a dataset. Do you know what to expect?

Recall: $Y, \hat{\beta}$ are R.V's since depend on the data.

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$$



$\hat{\beta} \sim N(\beta_{true}, ?)$

i.e) $E[\hat{\beta}] = \beta_{true}$ (i.e $\hat{\beta}$ is unbiased)

In practice it means that if we had M samples

$$\{x_1, \ldots, x_N\}_1 \to \hat{\beta}_1$$
$$\{x_1, \ldots, x_N\}_2 \to \hat{\beta}_2$$
$$\vdots$$
$$\{x_1, \ldots, x_N\}_M$$

we would see that

$$\hat{E}(\hat{\beta}) = \frac{1}{M} \sum_{j=1}^{m} \hat{\beta}_j \approx \beta_{true}$$

$$\left(\hat{?}\right)^{1/2} = \left(\frac{1}{M-1} \sum_{j=1}^{m} \left(\hat{\beta}_j - \hat{E}(\hat{\beta})\right)^2\right)^{1/2}$$

→ Estimate of the standard error of $\hat{\beta}$ (i.e. its standard deviation)