

Review

RV is function $\Omega \rightarrow \mathbb{R}$

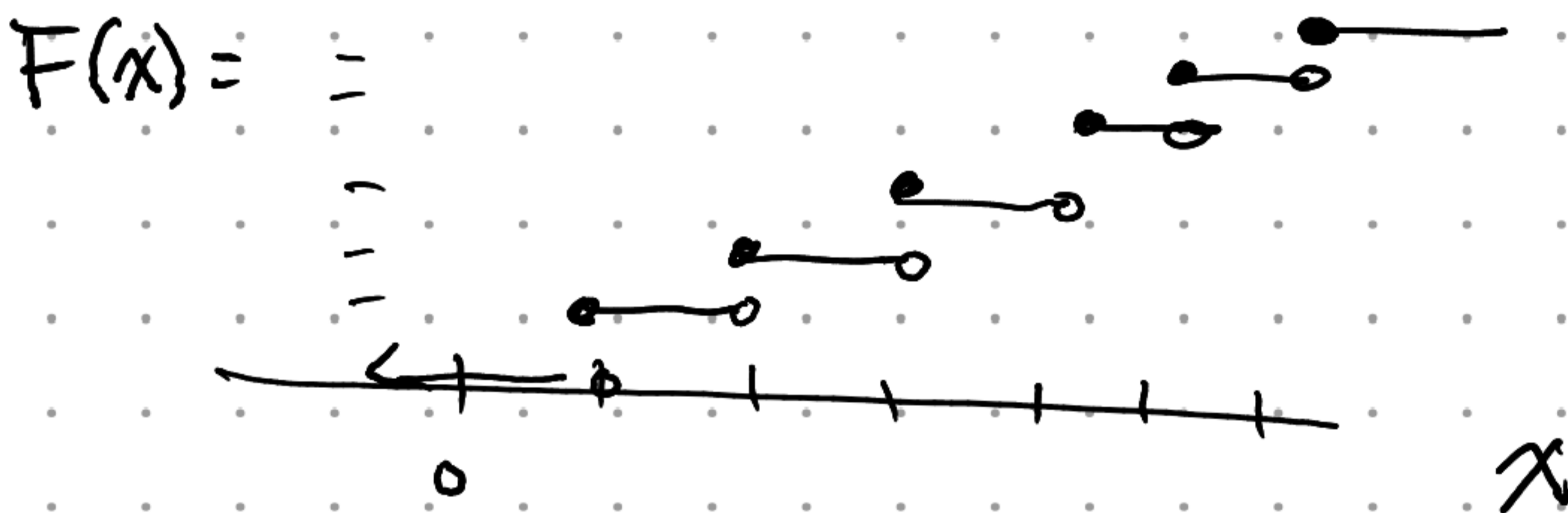
Distribution of RV X is described by

$$F(x) = P[X \leq x]$$

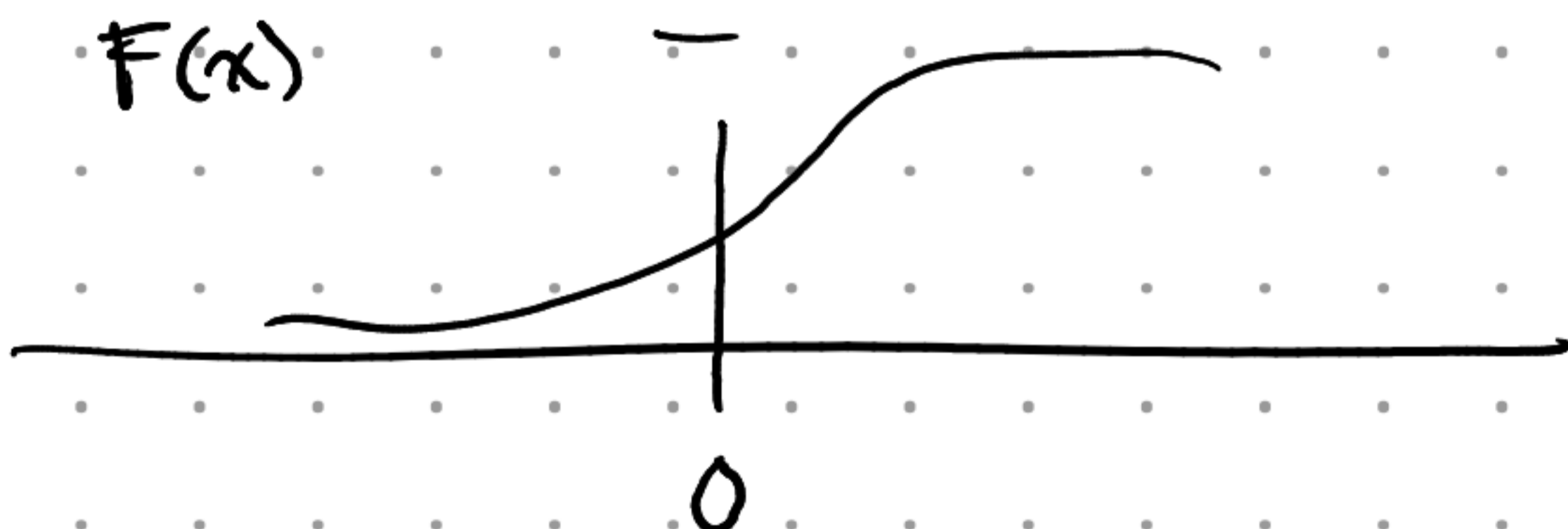
↑
local var

$$\left[\begin{array}{l} F(t) = P[X \leq t] \\ F_X(x) = P[X \leq x] \\ F_X(s) = P[X \leq s] \end{array} \right]$$

Ex. $X = \sum_{i=1}^n \dots$ w.p. $1/6$



Ex. X standard normal



In general F a CDF if

- $F(b) \geq F(a)$ if $b \geq a$

- $\lim_{x \rightarrow \infty} F(x) = 1$ $\lim_{x \rightarrow -\infty} F(x) = 0$

$$P[X \in (a, b]] = F(b) - F(a)$$

If X is discrete (values X can take are discrete)

pmf: $f(x) = P[X=x] = \lim_{c \rightarrow x} F(x) - F(c)$

If X is continuous, $P[X=x] = 0 \forall x$.

pdf: $f(x) = F'(x) = \lim_{dx \rightarrow 0} \frac{F(x+dx) - F(x)}{dx}$

$$= \lim_{dx \rightarrow 0} \frac{P[X \in (x, x+dx)]}{dx}$$

$$E[n(X)] = \int n(x) dF(x) = \begin{cases} \int n(x) f(x) dx & \begin{matrix} \swarrow \text{pdf} \\ X \text{ continuous} \end{matrix} \\ \sum_x n(x) f(x) & \begin{matrix} \swarrow \text{pmf} \\ X \text{ discrete} \end{matrix} \end{cases}$$

$$\text{Ex. } F(x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \leq x < 2 \\ 2/6 & 2 \leq x < 3 \\ 3/6 & 3 \leq x < 4 \\ 4/6 & 4 \leq x < 5 \\ 5/6 & 5 \leq x < 6 \\ 1 & 6 \leq x \end{cases}$$

$$\text{pmf } f(x) = \begin{cases} 1/6 & x=1 \\ 1/6 & x=2 \\ \vdots \\ 1/6 & x=6 \end{cases}$$

$$\int x dF(x) = \sum_x x f(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{7}{2}$$

$$\int (x - \mu)^2 dF(x), \quad \mu = \int x dF(x)$$

$$\begin{aligned} \rightarrow &= \sum_x (x - 3.5)^2 f(x) = (1 - 3.5)^2 \cdot \frac{1}{6} + \dots \\ &= \frac{3.5}{12} \end{aligned}$$

Statistics

Given data X_1, \dots, X_n , how do we learn something about the source of the data?

No assumptions = ill-posed

Ex. 1, 0, 1, 1, 1, 1, 0, 1. Where did this come from?

A statistical model is a set of distributions

$$\mathcal{F} = \{ F_\theta : \theta \in \Theta \}, \quad \Theta \subseteq \mathbb{R}^k$$

- parametric : $k < \infty$ $\theta = (\theta_1, \theta_2, \dots, \theta_k)$

- non-parametric : $k = \infty$

Ex $\theta = \{ \theta(x) : x \in \mathbb{R} \}$ $F_\theta(x) = \theta(x)$

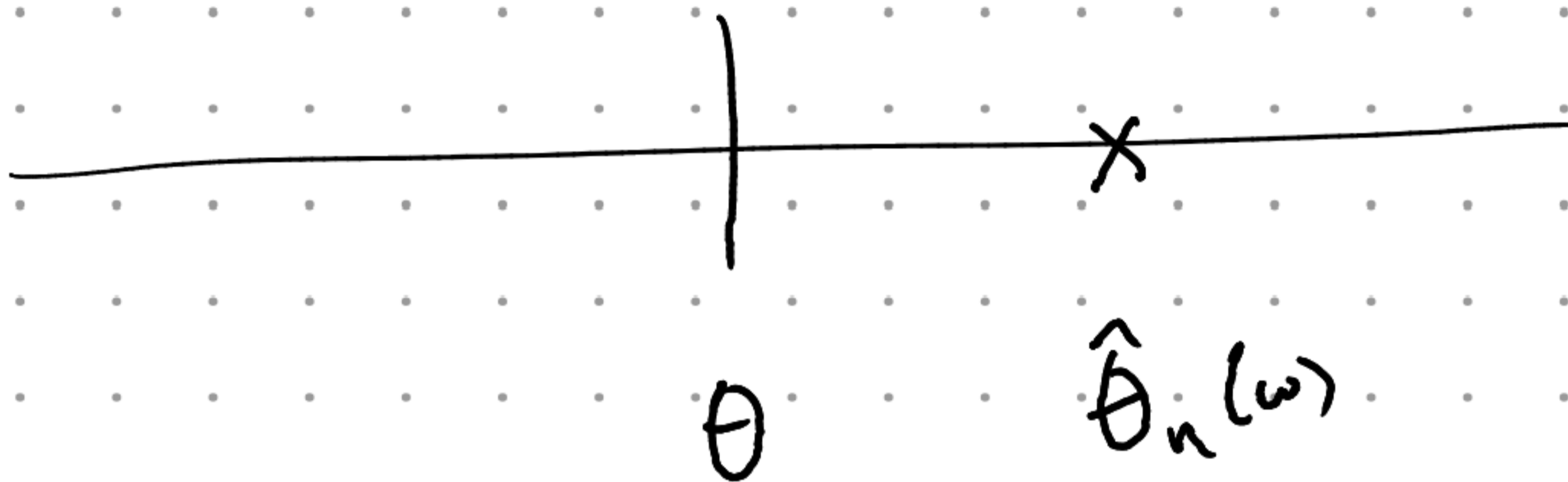
Assume $X_1, X_2, \dots, X_n \sim F_\theta$ (iid) for some $\theta \in \Theta$

How well can we learn something about θ/F_θ ?

- Here how-well means statistically (i.e. if we repeated this many times)

Suppose we use our data to get an estimate $\hat{\theta}_n$ for θ . How good is this estimate?

- Again, we can only think about this problem if $\hat{\theta}_n$ is a RV. For any fixed sample of our data, this is not meaningful.



- is this good or bad?
- Did we get lucky?

