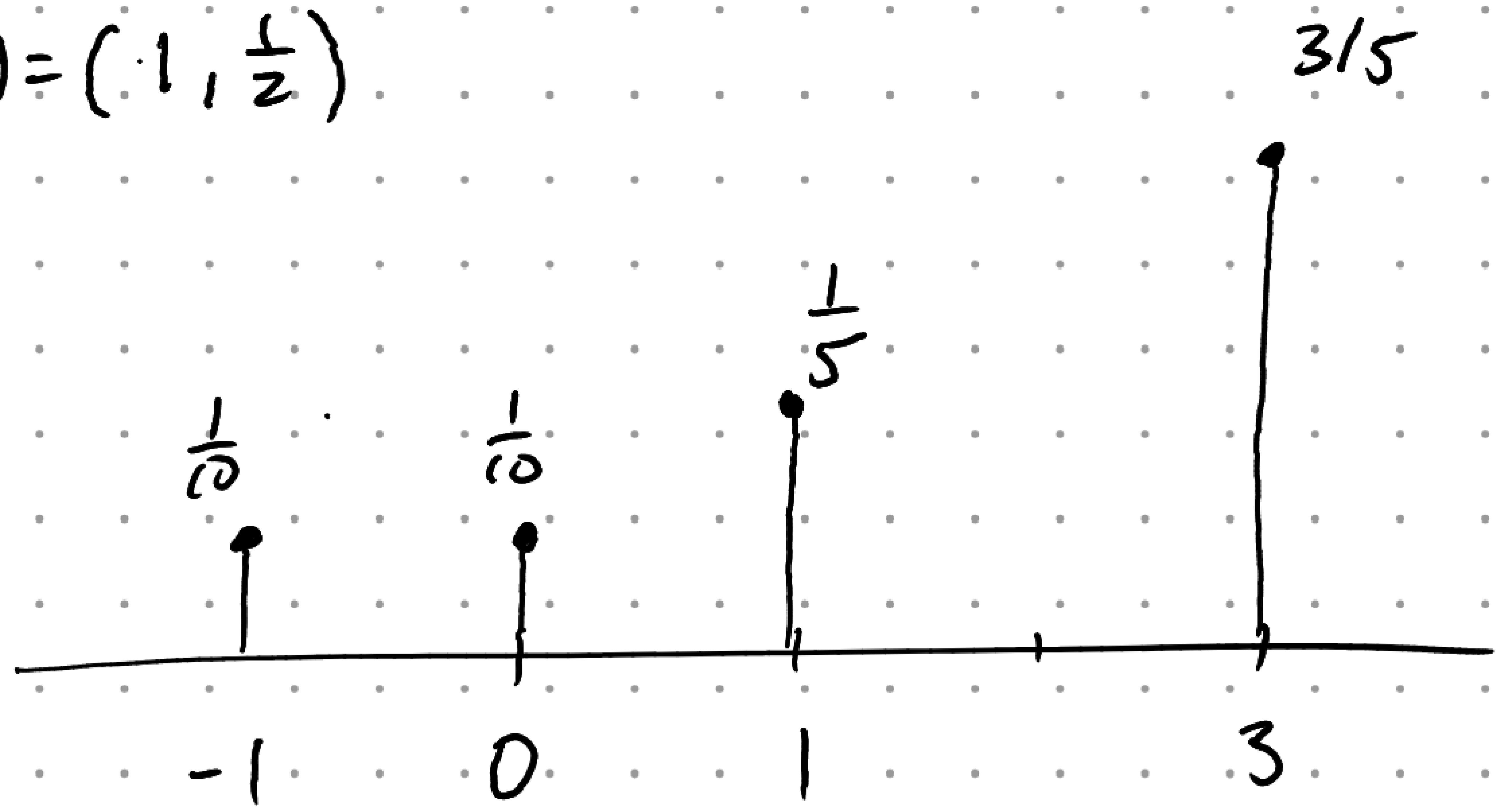
Suppose $f_\theta(x) = \begin{cases} \theta_1/10 & x = -1 \\ \theta_2/5 & x = 0 \\ \theta_1/5 & x = 1 \\ 1 - \left(\frac{\theta_1}{10} + \frac{\theta_2}{5} + \frac{\theta_1}{5}\right) & x = 3 \end{cases}$

$\downarrow$

$\theta = (\theta_1, \theta_2)$

If $X \sim F_\theta$, $\mathbb{P}[X = x] = f_\theta(x)$

$\theta = (1, \frac{1}{2})$



Sample $X_1, X_2, \ldots, X_n \sim F_\theta$ iid

- $\mathbb{P}[X_1 = -1, X_2 = 1, X_3 = 3]$ if $\theta = (1, \frac{1}{2})$
- $\mathbb{P}[X_1 = x_1, X_2 = x_2, X_3 = x_3]$ if $\theta = (\theta_1, \theta_2)$

$X_1, \dots, X_n$ data (for now, ignore how data was generated)

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} f_\theta(X_i)$$

This is a "local variable"

Suppose we now sample

$X_1', \dots, X_n' \sim F_\theta$ independently of $X_1, \dots, X_n$.

What is $\mathbb{P}[X_1' = X_1, \dots, X_n' = X_n \mid X_1, X_2, \dots, X_n]$?

$- \mathcal{L}_n(\theta)$

Idea: Let's choose $\hat{\theta}$ as the parameter that maximizes $\mathcal{L}_n(\theta)$ given the data $X_1, \dots, X_n$ we observed.

<u>Note</u> None of this requires assumption $X_1, \dots, X_n \sim F_\theta$. We use this assumption to understand how good our approach is.

<u>Properties</u>

Assume $X_1, \dots, X_n \sim F_\theta$. Then, the MLE estimator $\hat{\theta}_n$ is (under conditions)

- consistent: $\hat{\theta}_n \xrightarrow{P} \theta$
- equivariant: $\hat{\theta}_n$ MLE of θ, $g(\hat{\theta}_n)$ MLE of $g(\theta)$
- Asymptotically normal: $\dfrac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow{d} N(0,1)$
- Asymptotically optimal: as good as you can do

Ex. $X_1, \dots, X_n \sim N(\mu, \sigma^2) \qquad (\theta = (\mu, \sigma))$

$$\mathcal{L}_n(\mu,\sigma) = c \cdot \prod_{i=1}^{n} \frac{1}{\sigma} \exp\left(-\frac{(X_i-\mu)^2}{2\sigma^2}\right)$$

$$\ell_n(\mu,\sigma) = \log(c) + \sum_{i=1}^{n} \log\left(\frac{1}{\sigma}\right) - \frac{(X_i-\mu)^2}{2\sigma^2}$$

$$= \log(c) + n\log\left(\frac{1}{\sigma}\right) - \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial \ell_n(\mu,\sigma)}{\partial \mu} = \sum_{i=1}^{n} \frac{(X_i-\mu)}{2\sigma^2} = \frac{n(\bar{X}_n-\mu)}{2\sigma^2}$$

set equal to zero to find max

$$\frac{n(\bar{X}-\mu)}{2\sigma^2} = 0 \Longrightarrow \boxed{\mu = \bar{X}_n}$$

$$\frac{\partial \ell_n(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^3} = -\frac{n}{\sigma} + \frac{nS^2}{\sigma^3} + \frac{n(\bar{X}_n-\mu)^2}{\sigma^2}$$

where $S^2 = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

$$-\frac{n}{\sigma} + \frac{nS^2}{\sigma^3} + \cancelto{0}{\frac{n(\bar{X}-\mu)^2}{\sigma^2}} = 0 \Rightarrow \sigma^2 = S^2 \Rightarrow \boxed{\sigma = S}$$