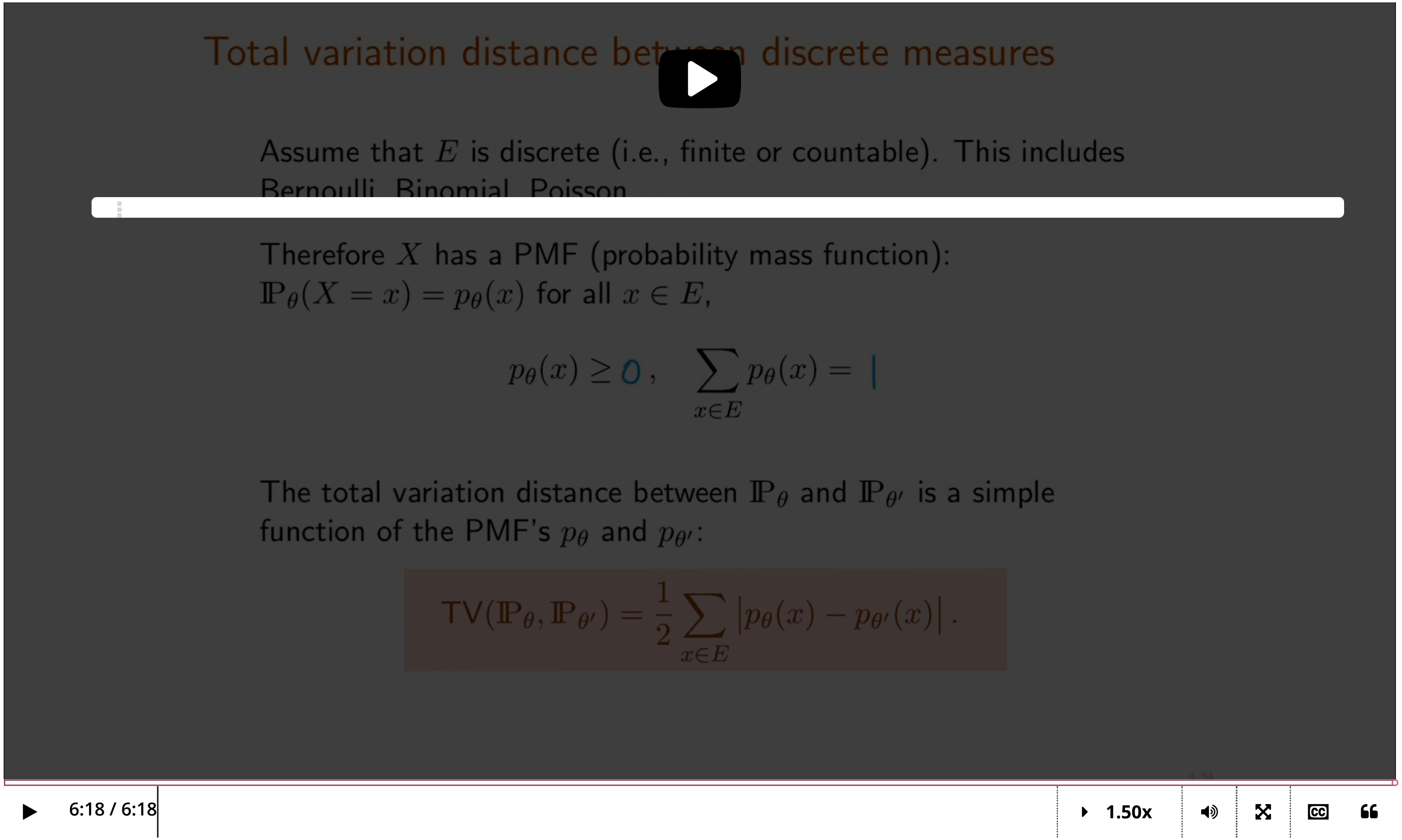
# 4. Introduction to Total Variation Distance

## Definition of Total Variation Distance




**Video**

Download video file

**Transcripts**

Download SubRip (.srt) file

Download Text (.txt) file

## Interpreting Total Variation Distance

1/1 point (graded)

Recall from lecture that the **total variation distance** between two probability measures $\mathbf{P}_\theta$ and $\mathbf{P}_{\theta'}$ with sample space $E$ is defined by

$$\mathrm{TV}\left(\mathbf{P}_\theta, \mathbf{P}_{\theta'}\right) = \max_{A \subset E} \left| \mathbf{P}_\theta(A) - \mathbf{P}_{\theta'}(A) \right|$$

Let $X_1, \ldots, X_n \overset{iid}{\sim} \mathbf{P}_{\theta^*}$ where $\theta^* \in \mathbb{R}$ is an unknown parameter. You construct a statistical model $(E, \{\mathbf{P}_\theta\}_{\theta \in \mathbb{R}})$ for your data. By analyzing your data, you are able to produce an estimator $\hat{\theta}$ such that the distributions $\mathbf{P}_{\hat{\theta}}$ and $\mathbf{P}_{\theta^*}$ are close in **total variation distance**. More precisely, you know that

$$\mathrm{TV}\left(\mathbf{P}_{\hat{\theta}}, \mathbf{P}_{\theta^*}\right) \leq \epsilon,$$

where $\epsilon$ is a very small positive number.

Which of the following can you conclude about the distributions $\mathbf{P}_{\hat{\theta}}$ and $\mathbf{P}_{\theta^*}$? (Choose all that apply.)

- [x] Let $A$ be an event. Then $\left|\mathbf{P}_{\theta^*}(A) - \mathbf{P}_{\hat{\theta}}(A)\right| \leq \epsilon$
- [x] Let $X \sim \mathbf{P}_{\theta^*}$, let $Y \sim \mathbf{P}_{\hat{\theta}}$, and suppose $a, b \in \mathbb{R}$ where $a \leq b$. Then $\left|\mathbf{P}_{\theta^*}(a \leq X \leq b) - \mathbf{P}_{\hat{\theta}}(a \leq Y \leq b)\right| \leq \epsilon$
- [ ] $|\theta^* - \hat{\theta}| \leq \epsilon$.

✔

**Solution:**

Recall that by definition,

$$\mathrm{TV}\left(\mathbf{P}_{\hat{\theta}}, \mathbf{P}_{\theta^*}\right) = \max_{A \subset E} \left| \mathbf{P}_{\hat{\theta}}(A) - \mathbf{P}_{\theta^*}(A) \right|$$

where the maximum is over all events $A$. Since we are given that $\mathrm{TV}\left(\mathbf{P}_{\hat{\theta}}, \mathbf{P}_{\theta^*}\right) \leq \epsilon$, we conclude that $\left|\mathbf{P}_{\hat{\theta}}(A) - \mathbf{P}_{\theta^*}(A)\right| \leq \epsilon$ for every event $A$. Hence, the first choice is correct.

Let $A$ be the event given by the interval $(a, b)$. Then,

$$\left|\mathbf{P}_{\theta^*}(a \leq X \leq b) - \mathbf{P}_{\hat{\theta}}(a \leq Y \leq b)\right| \leq \epsilon$$

is the same as saying $\left|\mathbf{P}_{\hat{\theta}}(A) - \mathbf{P}_{\theta^*}(A)\right| \leq \epsilon$. Thus, the second choice is true as well.

The third choice, "$|\theta^* - \hat{\theta}| \leq \epsilon$.", is incorrect. In general, even if distributions $\mathbf{P}_{\hat{\theta}}$ and $\mathbf{P}_{\theta^*}$ are close, there is no reason to expect the parameters $\theta^*$ and $\hat{\theta}$ to be close. To conclude that the estimated parameter is close to the true parameter given their distributions are close, we would need some assumptions on the map $\theta \mapsto \mathbf{P}_{\theta}$. No such assumption is given here.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

## Discussion

Hide Discussion

**Topic:** Unit 3 Methods of Estimation:Lecture 8: Distance measures between distributions / 4. Introduction to Total Variation Distance

Add a Post

Show all posts ▼ by recent activity ▼

There are no posts in this topic yet.

✖