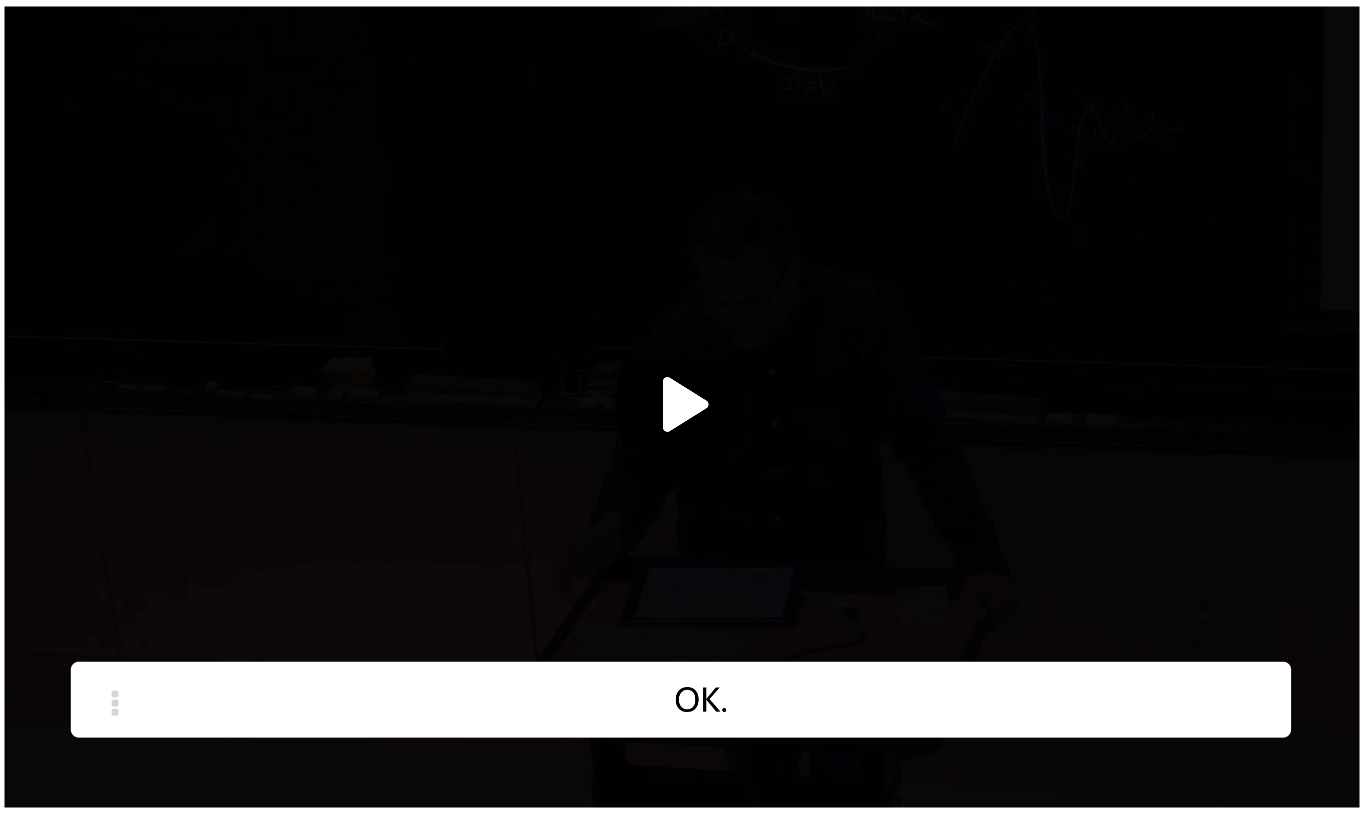
7. Linear Regression - Basic Setup

Linear Regression: The Function for Conditional Expectation of Y Given a value x



7:39 / 7:39 1.0x

is, make a transformation of your x

so that you actually only need to guess maybe

the tilt, or something like this with this curve.

So you can do data transformation

before you do linear regression, right?

You could not do y on to x, but y

on to some cosine of x for example, or x squared.

OK.

End of transcript. Skip to the start.

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

xuetangX.com 学堂在线

In **Linear Regression** , *we will work with the assumption that the regression function* $\nu(x) := \mathbb{E}[Y|X=x]$ *is linear*, so that

$$\nu(x) = a + bx$$

for some pair (a, b).

In this unit, we will be studying the **Least Squares Estimator** . It is an estimator $(\hat{a}, \hat{b})$ so that $\hat{Y} = \hat{a} + \hat{b}X$ is “close" (in some distance metric) to the actual Y as often as possible.

A Minimization Problem

1/1 point (graded)
Let X be an arbitrary random variable, with mean μ and variance σ^2. In terms of μ and σ^2, which scalar k is the unique minimizer of the function $f(k) = \mathbb{E}[(X-k)^2]$?

Hint: Write $f(k)$ *as a quadratic in* k.

mu ✔ **Answer:** mu

μ

STANDARD NOTATION

Solution:

Observe that we can re-write $f(k)$ as

$$\begin{aligned} f(k) &= \mathbb{E}\left[X^2 - 2kX + k^2\right] \\ &= k^2 - 2\mathbb{E}[X]\,k + \mathbb{E}\left[X^2\right] \end{aligned}$$

This is a positive quadratic in k, with a unique minimizer at the center of the parabola. To be thorough, we will take the derivative with respect to k:

$$f'(k) = 2k - 2\mathbb{E}[X]$$

Setting this equal to zero gives $k = \mathbb{E}[X] = \mu$. Observe that the answer does not depend on σ.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

An Estimator

1/1 point (graded)
Let (X, Y) be a pair of random variables for which the regression function $\nu(x) = \mathbb{E}[Y|X = x]$ takes the form

$$\nu(x) = a + bx$$

for some pair of real numbers (a, b).

What is a random variable $\hat{Y}$ that is a function of X that minimizes

$$\mathbb{E}_{X,Y}\left[(Y - \hat{Y})^2\right]$$

over all possible choices of $\hat{Y}$? Enter your answer in terms of a, b and the random variable X (capital letter "X").

(Remark: for a clean, quick solution, it may be helpful to review the law of iterated expectations: $\mathbb{E}_{X,Y}[\cdot] = \mathbb{E}_X\left[\mathbb{E}_Y[\cdot \mid X]\right]$, where $\mathbb{E}_Y[\cdot|X]$ denotes the conditional expectation, which is a random variable. Use the insight from the previous exercise.)

a+b*X ✔ **Answer:** a + b*X

$a + b \cdot X$

STANDARD NOTATION

Solution:

By the law of iterated expectations, we may rewrite the provided expectation as

$$\mathbb{E}_X\left[\mathbb{E}_Y\left[(Y - \hat{Y})^2 | X\right]\right].$$

For each realization x of X, the previous exercise tells us that $\mathbb{E}_Y\left[(Y - \hat{y})^2 | X = x\right]$ is minimized by $\hat{y} = \nu(x)$. Since $\nu(x) = a + bx$ is a minimizer for each choice of x, $\nu(X)$ is a minimizer over all choices of $\hat{Y}$.

These two exercises verify that the Least Squares Estimator is consistent in the following sense: **using the actual distribution on (X, Y), the true pair (a, b) itself is a least squares estimator.** It may or may not be unique; we will address this in the following sections.