edX

Course > Unit 6 Linear Regression > Lectures 19: Linear Regression 1 > 6. Plots of Conditional Distributions and Conditional Quantiles and Box-and-Whisker Plots
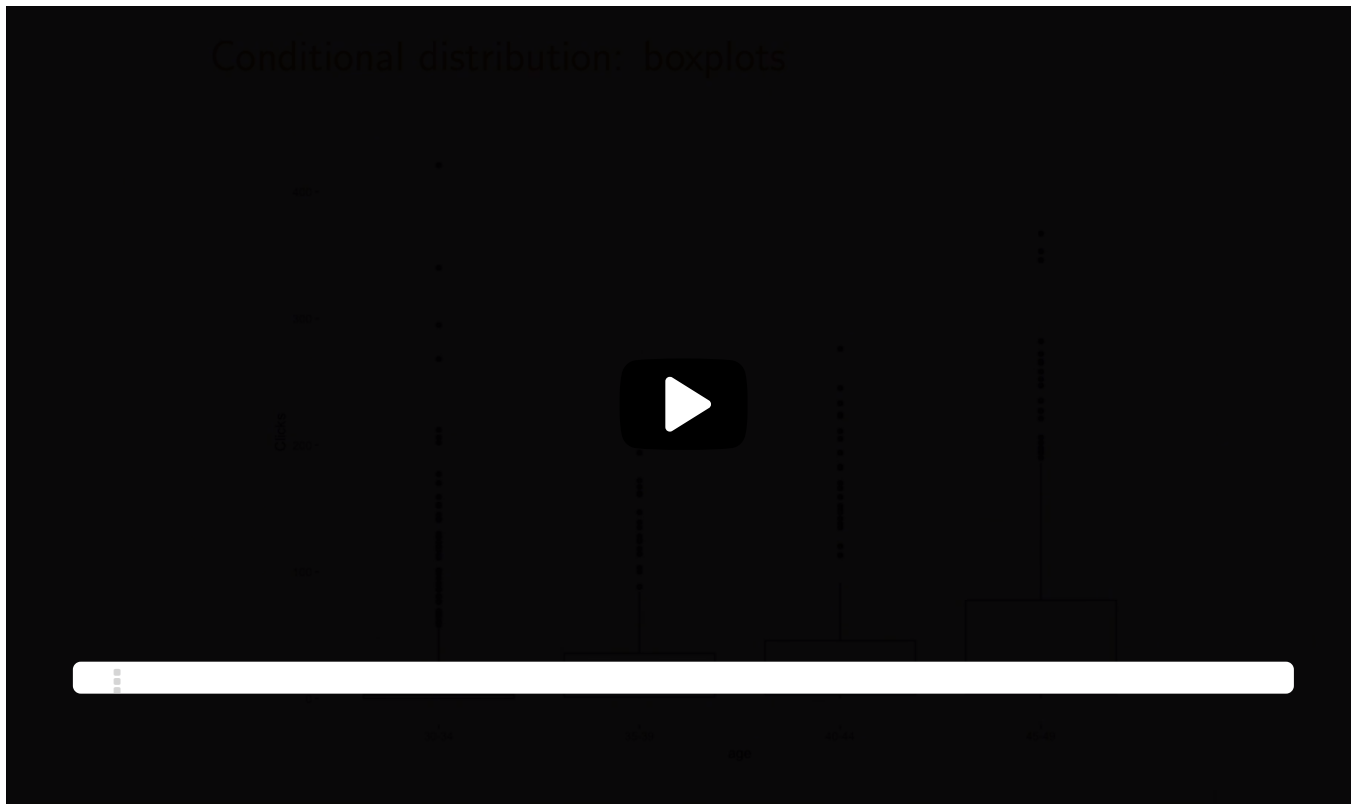
6. Plots of Conditional Distributions and Conditional Quantiles and Box-and-Whisker Plots

Plots of Conditional Distributions and Conditional Quantiles and Boxplots

group than there is for other age groups.

But you can see actually an upward trend, right?

OK.

So this is one way to represent conditional distributions.

But if I wanted to summarize it by one number in this case,

I would probably look at what?

Conditional median, right?

7:24 / 7:24   ▶ 1.0x ◀)) ⤢ [CC] "

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

The Shapes of Joint and Conditional Distributions

1/1 point (graded)

Let $f(x, y)$ be the joint pdf of the pair of **continuous** random variables (X, Y). Select from the following all statements that are correct.

☑ $f(y \mid x) = \frac{f(x,y)}{f(x)}$, provided $f(x) \neq 0$ ✔

☑ $f(y \mid x)$, for each x, is a scaled version of the slice of the 3-D plot of $f(x, y)$ across the x intercept. ✔

☐ $f(y|x)$ is the probability that $Y = y$ given $X = x$.

✔

Solution:

The first two choices are correct. The first choice is the definition of the conditional pdf. The second choice is correct because it follows from the definition that for a given x, $f(y|x)$ is equal to the slice of $f(x, y)$ at the x intercept scaled by $\frac{1}{f(x)}$.

The third choice is not correct because, among many things, $f(y|x)$ can take values greater than 1 and still be a valid pdf and a probability can never be greater than 1. Further, for a continuous random variable with a continuous pdf the probability that the random variable takes on any particular value is equal to 0.

ℹ Answers are displayed within the problem

Box-and-Whisker Plot

1/1 point (graded)

The following is a box-and-whisker plot drawn using 100 y axis points for each x value $1, 2, 3, 4$:
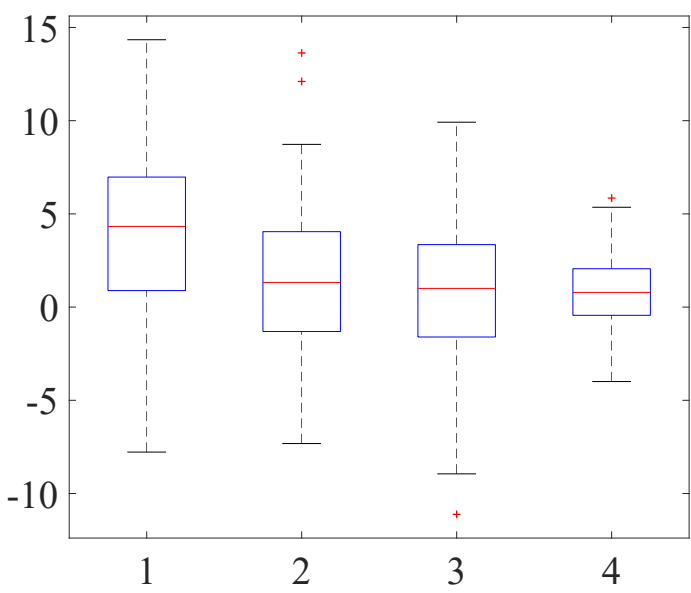
Which of the following statements accurately reflect the box-and-whisker plot shown above?

☐ The y values have a similar standard deviation for all values of $x = 1, 2, 3, 4$.

☑ The middle 50 percentile of y values is more concentrated for $x = 4$ when compared to the other three x values. ✔

☑ The conditional distribution of y given x does not appear to be heavily skewed one way or another for all values of x. ✔

☑ If we assume as ground truth that for every x the conditional distribution of y given x is symmetric, then it is reasonable to conclude that the mean of the conditional distributions is non-increasing with x. ✔

✔

Solution:

The second, third, and fourth statements reflect what we see in the box-and-whisker plot.

ℹ Answers are displayed within the problem

Concept Check: Estimation for Each x

1/1 point (graded)

Is the following statement true or false?

"For a given x, upon observing samples $Y_1, \ldots, Y_n$ we have seen an estimation technique in this class to estimate the mean, median, and quantiles of Y for this x even in the absence of a statistical model for Y given this x."

⦿ True ✔

○ False

Solution:

The answer is **True**. M-estimation can be used to estimate the mean, median, and quantiles of Y for a given x even without having to assume a statistical model for the same.

Submit   You have used 1 of 1 attempt

ℹ Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 6 Linear Regression:Lectures 19: Linear Regression 1 / 6. Plots of Conditional Distributions and Conditional Quantiles and Box-and-Whisker Plots