edX

Course   **Discussion**   Progress   Resources

◀   ▶

Discussion

ya_mukhin   *Moderator, Student*

4 discussions started   430 comments

Show all posts ▼	by recent activity ▼
💬 **[STAFF] Solution for "Statistics for LSE" is missing some terms.** u^t on RHS of equation on line 2. Extra X^T on LHS two times in each parentheses in equation on line 3.	7
? **[STAFF] Theoretical Linear Regression Visualized II: shouldn't the E[Y] > E[Y'] be strict inequality?** Otherwise it seems like there is some ambiguity in answer choices.	5
? **Why the Hypothesis that the Mean for x=3 is less than Mean for x=4 is rejected?** Few months we were doing Hypothesis testing. Now you ask question: *"If we assume as ground truth that for every x the conditional distribution of y given x is symmetric, then is it reas...	4
? **How to check I start the 48 hours timed exam or not.** Dear TA, I opened and read this rule page already. I am a bit afraid that my time starts clouting down,,, How to check I really kicked off the exam?	53
💬 **Kaggle facebook data** The data set used in the video can be found here https://www.kaggle.com/chrisbow/an-introduction-to-facebook-ad-analysis-using-r If anyone can show me how to write the R/ggplot() c...	7
💬 **No timer!** I came here and found no timer, but having seen the questions, thought I ought to start. My start time was November 17, 17:52 GMT.	8
? **[Staff] No timer.** I didn't see any countdown timer as mid - term 1. Is there anything wrong ?	3

? Prediction error
2
Does it increase with more data? And decreases if there are more parameters to estimate, given the same data?

? [Staff] The LSE problem has two separate definitions of p
2

? Slides for Lec 19 & 20: Linear Regression
3
Where are the slides used in those lectures posted? I cant find it under the Resource Tab. Thanks!

💬 [STAFF] When the exam ends...
2
...Please also add the explanations for all of the answers. There are topics that I don't know how well I understood them, and explanations would be quite useful (e.g. the final exam).

💬 [STAFF] Missing Format Preview for Upper Bound of Confidence Interval
3
It's probably trivial, but it's also true. Just figured I'd point it out for future iterations of the course/consistency purposes.

? [Staff]: you have an error in your sigma squared subscripts in the last question

---

All Posts

Kaggle facebook data

[+]

discussion posted 6 days ago by **mbh038**

[★]

The data set used in the video can be found here https://www.kaggle.com/chrisbow/an-introduction-to-facebook-ad-analysis-using-r

[...]

If anyone can show me how to write the R/ggplot() code to do the density plot shown at around 3:00 I would be very grateful!

This post is visible to everyone.

Add a Response   2 responses

Erocha   [+]
6 days ago   [...]

Look for `geom_density` .

Thanks, I know to use `geom_density` but how to get the precise form of plot that we see in the video escapes me, despite looking everywhere I can think of, and trying various ideas.

I've got this, which doesn't work, but can't work out how to modify it such that it does

```
g<-ggplot(fb,aes(x=age,fill=Clicks)+
geom_density(position='fill')
g
```

the data are in a data frame (tibble) called `fb` :

```
glimpse(fb)
Observations: 1,143
Variables: 11

$ age   <chr> "30-34", "30-34", "30-34", "30-34", "30-34", "30-34", "30-34", "30-34", "30-3…

$ Clicks  <dbl> 1, 2, 0, 1, 1, 0, 3, 1, 1, 3, 0, 0, 0, 0, 7, 0, 1, 0, 1, 4, 0, 2, 2, 2, 0, 0,…
```

I wonder if I need to somehow use `group_by()` .. to get a density plot for each age.

posted 6 days ago by **mbh038**

Add a comment

KYPark

a day ago

You can use ggridges package.

example>

library(ggplot2)

library(ggridges)

x = rep(c("a", "b"), each = 100)
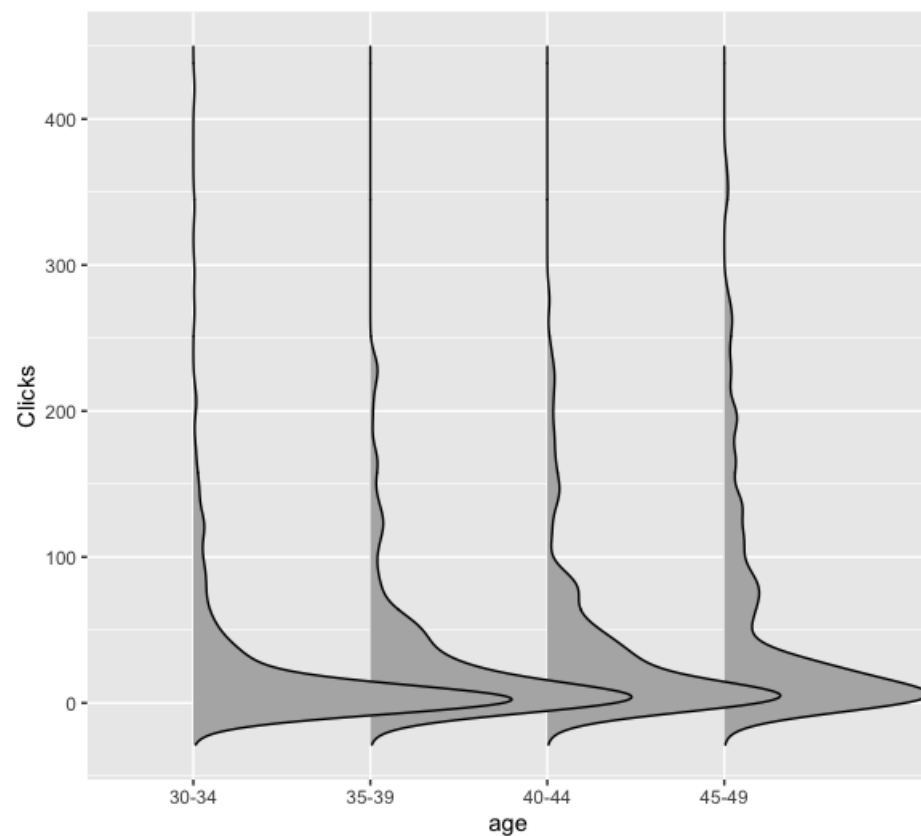
y = c(rnorm(100, 0, 1), rnorm(100, 1, 2))

db = data.frame(x, y)

ggplot(db, aes(y, x)) + geom_density_ridges(aes(fill = x)) + coord_flip()

That's it, thanks very much!

```
library(tidyverse)
library(ggridges)
fb<-read_csv('KAG_conversion_data.csv',skip=1)
g2<-ggplot(fb,aes(Clicks,age))+
  geom_density_ridges(aes(fill=Clicks))+
  coord_flip()
g2
```


Though weirdly, as with the Prof's, the density plots extend to negative click values.

You have to tilt your head and maybe quint a bit to see hints in this of a Joy Division album cover or even of a signal of pulsars.

See Pop Culture Pulsars

posted a day ago by **mbh038**

Hi @mbh038, thank you for sharing your code for the density plots! The reason for seeing positive density over negative values is because the nonparametric density estimator adds a small Gaussian distribution around every observation and then sums up all the small Gaussians into a mixture distribution with a density that you see on the plot. Because there are many observations at the origin, and because the Gaussian puts mass above and below the observation, you see positive density over negative values.

posted a day ago by **ya_mukhin**

Ah! I see. Thanks.

posted about 13 hours ago by **mbh038**

Add a comment

Showing all responses

Add a response:

Preview

Submit

---

edX

About

edX for Business

Legal

Terms of Service & Honor Code

Privacy Policy

Connect

Blog

Contact Us

Accessibility Policy

Help Center