

<u>Help</u>

HuitianDiao 🗸

<u>Course</u> <u>Progress</u> <u>Dates</u> <u>Discussion</u> <u>Resources</u>

★ Course / Unit 4 Hypot... / Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov t...

()



## 8. Kolmogorov-Smirnov Test Statistic Pivotal Under Null

□ Bookmark this page

## Non-asymptotic Distribution, Generating Data from a Given Distribution

Start of transcript. Skip to the end.



So It turns out that this result is also true.
So what did we want?
Why did we use the Donsker's theorem?
Because Donsker's theorem tells us
that no matter what the true f is,
asymptotically,
or even what f0 is, asymptotically,
square root of n, fn minus f not, if f not
equal to f,

▶ 1.50x

will converge to compthing which is

Video

Download video file

0:00 / 0:00

**Transcripts** 

X

<u>Download SubRip (.srt) file</u> <u>Download Text (.txt) file</u>

CC

66

### CDF as a Random Function

3 points possible (graded)

Let X be a random varible with invertible cdf  $F_X$ . Define another random variable  $Y=F_X\left(X\right)$ . Find the cdf  $F_Y$  of Y .

For t < 0:

$$F_{Y}\left( t
ight) =% {\displaystyle\int\limits_{t=0}^{\infty }} {\displaystyle$$

For  $t \geq 1$ :

$$F_{Y}\left( t
ight) =% {\displaystyle\int\limits_{t=0}^{\infty }} {\displaystyle$$

For  $0 \leq t < 1$ :

$$F_{Y}\left( t
ight) =% {\displaystyle\int\limits_{t=0}^{\infty }} {\displaystyle$$

(What is the distribution of Y?)

STANDARD NOTATION

Submit

You have used 0 of 3 attempts

omogorov-3mmov rest statistic as a Pivotai Distribution onder Num hypothesis

Let  $X_1,\ldots,X_n$  be iid samples with unknown cdf  $F_X$ . For simplicity, restrict to the cases when  $F_X$  is invertible.

Recall the goal of the Kolmogorov-Smirnov Test goodness of fit test is to decide between the hypotheses

$$H_0$$
 :  $F_X = F^0$ 

$$H_1 : F_X \neq F^0.$$

Recall also the Kolmogorov-Smirnov test statistic:

$$T_{n} \ = \ \sqrt{n} \sup_{t \in \mathbb{R}} \left| F_{n} \left( t 
ight) - F^{0} \left( t 
ight) 
ight|$$

Assuming  $H_0$  is true, then  $T_n$  becomes

$$T_{n} \, = \, \sqrt{n} \sup_{t \in \mathbb{R}} \left| F_{n} \left( t 
ight) - F_{X} \left( t 
ight) 
ight|$$

We will see that under the null hypothesis, the distribution of  $T_n$  does not depend on the distribution of the data  $X_i$ , i.e.  $T_n$  is pivotal, and this is true for any n, not only for large n.

The trick is to make a change of variables. Let  $ilde{t}=F_{X}\left(t
ight)$  , then  $\,t=F_{X}^{-1}\left( ilde{t}\,
ight)$  . We have

$$egin{aligned} T_n &=& \sqrt{n} \sup_{t \in \mathbb{R}} \left| F_n\left(t
ight) - F_X\left(t
ight) 
ight| \ &=& \sqrt{n} \sup_{t \in \mathbb{R}} \left| \left(rac{1}{n} \sum_{i=1}^n \mathbf{1}\left(X_i \le t
ight)
ight) - F_X\left(t
ight) 
ight| \qquad ext{(definition of empirical cdf)} \ &=& \sqrt{n} \sup_{t \in \mathbb{R}} \left| \left(rac{1}{n} \sum_{i=1}^n \mathbf{1}\left(F_X\left(X_i
ight) \le F_X\left(t
ight)
ight) 
ight) - F_X\left(t
ight) 
ight| \qquad ext{(apply $F_X$ to both sides of inequality )} \ &=& \sqrt{n} \sup_{ ilde{t} \in (0,1)} \left| \left(rac{1}{n} \sum_{i=1}^n \mathbf{1}\left(Y_i \le ilde{t}
ight) 
ight) - ilde{t} 
ight| \qquad ext{where $Y_i \sim \mathsf{Unif}\left(0,1
ight)$.} \end{aligned}$$

Discussion

**Hide Discussion** 

**Topic:** Unit 4 Hypothesis testing:Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov test, Kolmogorov-Lilliefors test, Quantile-Quantile Plots / 8. Kolmogorov-Smirnov Test Statistic Pivotal Under Null

Add a Post

Show all posts 💙	by recent activity 🗸
Any hint for the excersice?	4

Previous

Next >



#### EUV

**About** 

<u>Affiliates</u>

edX for Business

Open edX

<u>Careers</u>

<u>News</u>

# Legal

Terms of Service & Honor Code

Privacy Policy

Accessibility Policy

**Trademark Policy** 

<u>Sitemap</u>

# **Connect**

<u>Blog</u>

Contact Us

Help Center

Media Kit

**Donate** 















© 2021 edX Inc. All rights reserved.

深圳市恒宇博科技有限公司 <u>粤ICP备17044299号-2</u>