Finding quantile numerically in R

Asked yesterday   Modified today   Viewed 62 times   ✪ Part of R Language Collective

▲

4

▼

🔖

🕘

I'm looking for an R code that would help me find the quantile when the pdf is given numerically. That is, say my data is

```
x =
 c(0.00,0.05,0.10,0.15,0.20,0.25,0.30,0.35,0.40,0.45,0.50,0.55,0.60,0.65,0.70,0.75,0.80,0.
```

and the corresponding pdf values are

```
pdf =
 c(0.000000000,1.221759375,1.968300000,2.349028125,2.457600000,2.373046875,2.160900000,1.8
```

I'd like to find the 95th quantile. I found an example here, however, the pdf is considered to be a function in this example, which is not my case. Thank you ahead for any help.

    r   statistics   linear-regression   numerical-methods   quantile   Edit tags

Share Edit Follow Close Flag

edited yesterday	asked yesterday
TT ThomasIsCoding **97.1k** 9 24 81	🔷 EM823823 **153** 5
✪ Recognized by R Language Collective	

2 Answers

Sorted by:
Reset to default

[Date modified (newest first) ⇕]

▲

0

▼

🔖

🕘

We can also do it in 2 steps:

- first approximate the (cumulative) distribution function `F_x` from the datapoints (with polynomial regression, e.g.)
- then use the definition of the quantile function (as inverse function of the CDF)

`F_X(x) = P(x >= x) = p` and `Q(p) = inv(F_x)(p)`

```
x =
 c(0.00,0.05,0.10,0.15,0.20,0.25,0.30,0.35,0.40,0.45,0.50,0.55,0.60,0.65,0.70,0.75,0.80,0.
f =
 c(0.000000000,1.221759375,1.968300000,2.349028125,2.457600000,2.373046875,2.160900000,1.8
F = cumsum(f) / sum(f) # normalize

k <- 5 # degree of the polynomial to approximate F_X(.) from the data
m <- lm(F ~ poly(x, degree=k))
```

```
F_x <- function(x) {     # CDF
  predict(m, newdata=data.frame(x=x))
}

inverse = function (f, lower = -Inf, upper = Inf) { # Q as inverse of CDF
  function (y) uniroot((function (x) f(x) - y), lower = lower, upper = upper)[1]$root
}

# now compute quantile
quantile = inverse(function(x) F_x(x), 0, 1)
quantile(0.95)
# [1] 0.5535394
```

The next figure shows how the 95% percentile value changes as the degree of polynomial k to approximate the CDF F_x(.) changes:
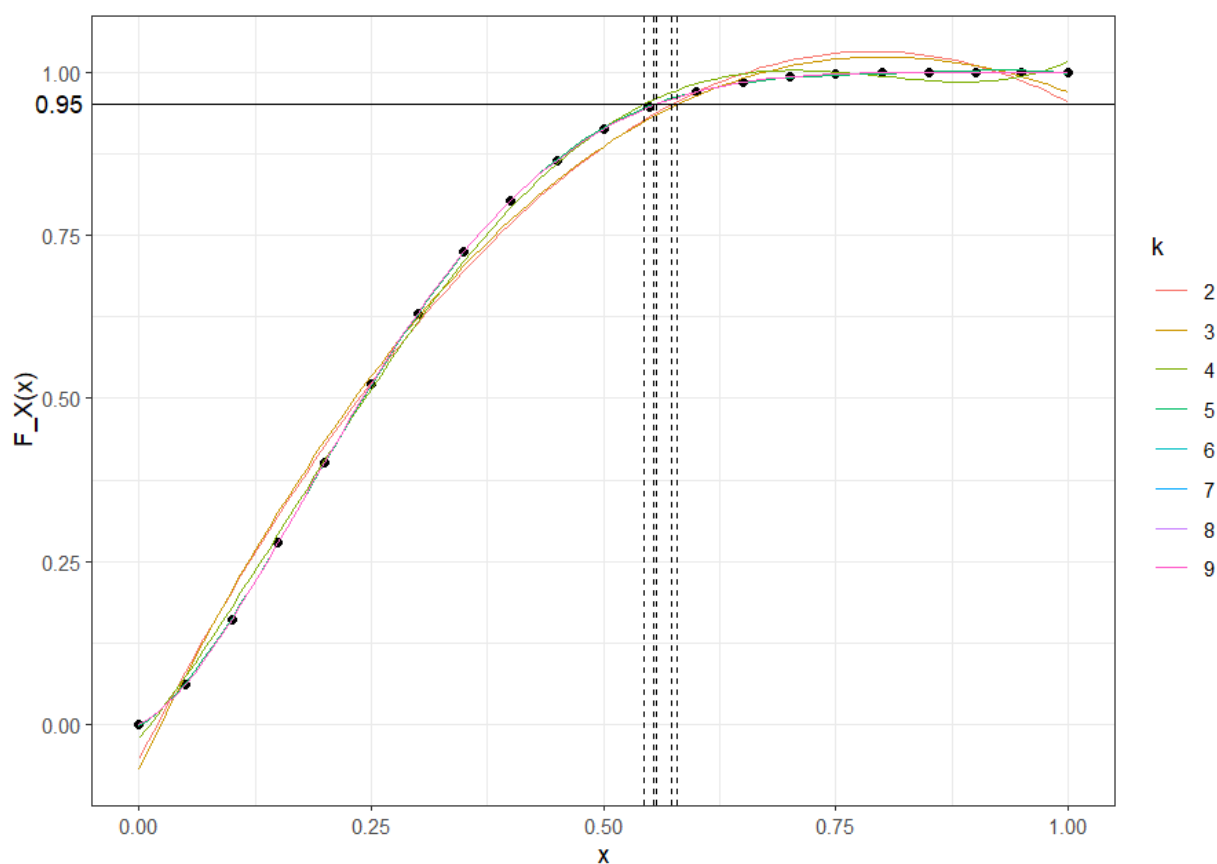
Share Edit Delete Flag

answered 3 mins ago

Sandipan Dey
21.7k 3 52 63

You can calculate an approximate 95%-quantile according to the definition of quantile, e.g.,

0

```
> max(x[cumsum(pdf) <= 0.95 * sum(pdf)])
[1] 0.55
```
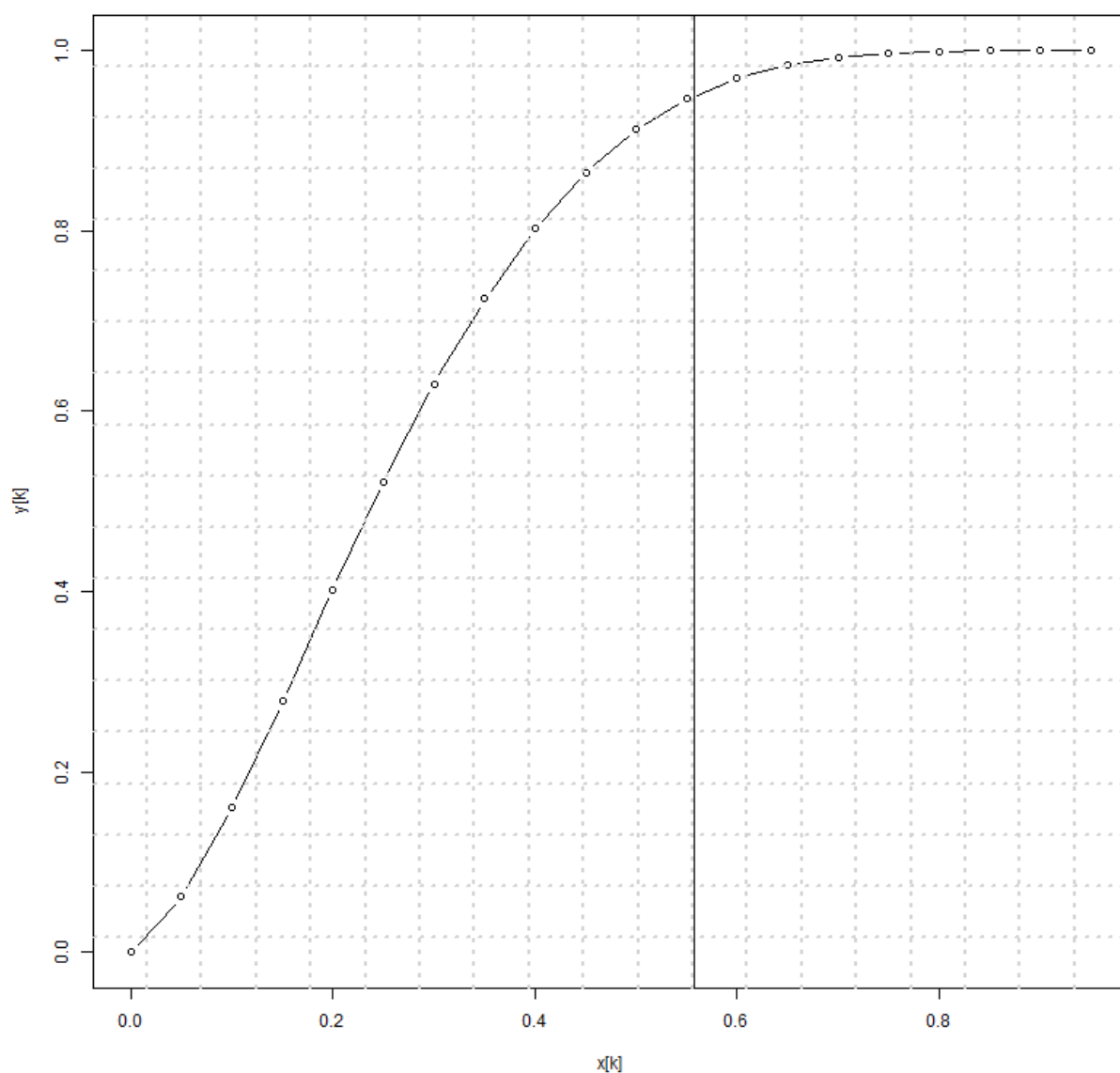
Or, you can try approxfun like below

```
> y <- cumsum(pdf) / sum(pdf)

> k <- !duplicated(y)

> approxfun(y[k],x[k])(0.95)
[1] 0.5577952
```

Below is the visualization

```
plot(x[k], y[k], type = "b")
abline(v = q)
grid(nx = 19, ny = 19, lwd = 2)
```


Share Edit Follow Flag   edited 20 hours ago   answered yesterday

ThomasIsCoding
97.1k 9   24   81

✿ Recognized by R Language
Collective

I believe this is what I was looking for. Thanks! – EM823823 8 hours ago