Test

Stat645

Confidence intervals for variance

Hadley Wickham

1. Test
2. Review
3. Case study: variance
4. Point estimate: MLE and properties
5. Cochran’s theorem
6. Interval estimate

Test

Review

Steps

Identify distribution that connects estimator and true value.

Form confidence interval for known (sampling) distribution.

Write as probability statement.

Back transform.

Write as interval.

Your turn

Recall the four distributions that connect possible estimators with true values that we learned about last time.

X_i iid, and n large:

$$\frac{\bar{X}_n - E(\bar{X})}{sd(\bar{X})} \dot{\sim} Z$$

and even more approximately

$$\frac{\bar{X}_n - E(\bar{X})}{\hat{sd}(\bar{X})} \dot{\sim} Z$$

$$X_i \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$$

$$\frac{\bar{X}_n - E(\bar{X})}{sd(\bar{X})} \sim Z$$

but

$$\frac{\bar{X}_n - E(\bar{X})}{\hat{sd}(\bar{X})} \sim t_{n-1}$$

Case study

$$X_i \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$

Unknown mean
Unknown variance

Variance

Develop a maximum likelihood estimator for the variance.

Look at its properties (bias, consistency).

See an alternative estimator and its properties.

Create a confidence interval for a non-symmetric distribution.

MLE Variance

Your turn

Recall the four steps for working out a maximum likelihood estimator.

Once you've recalled them, start working on the MLE for the variance.

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i}^{n}(X_i - \bar{X}_n)^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j$$

Properties

What's the expected value of our estimator?

Is it biased? Is it consistent?

If it is biased, how could you fix it?

Cochcran's theorem

Steps

Identify distribution that connects estimator and true value.

Form confidence interval for known (sampling) distribution.

Write as probability statement.

Back transform.

Write as interval.

$$X_i \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \boxed{\chi^2(n-1)}$$

New distribution!

Five fun facts about

$$\chi^2$$

Your turn

Fill in the blanks to prove:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Interval estimate

Steps

Identify distribution that connects estimator and true value.

Form confidence interval for known (sampling) distribution.

Write as probability statement.

Back transform.

Write as interval.

Interval for known dist

Want $P(a < Q < b) = 1 - \alpha$, and $b - a$ to be as small as possible.

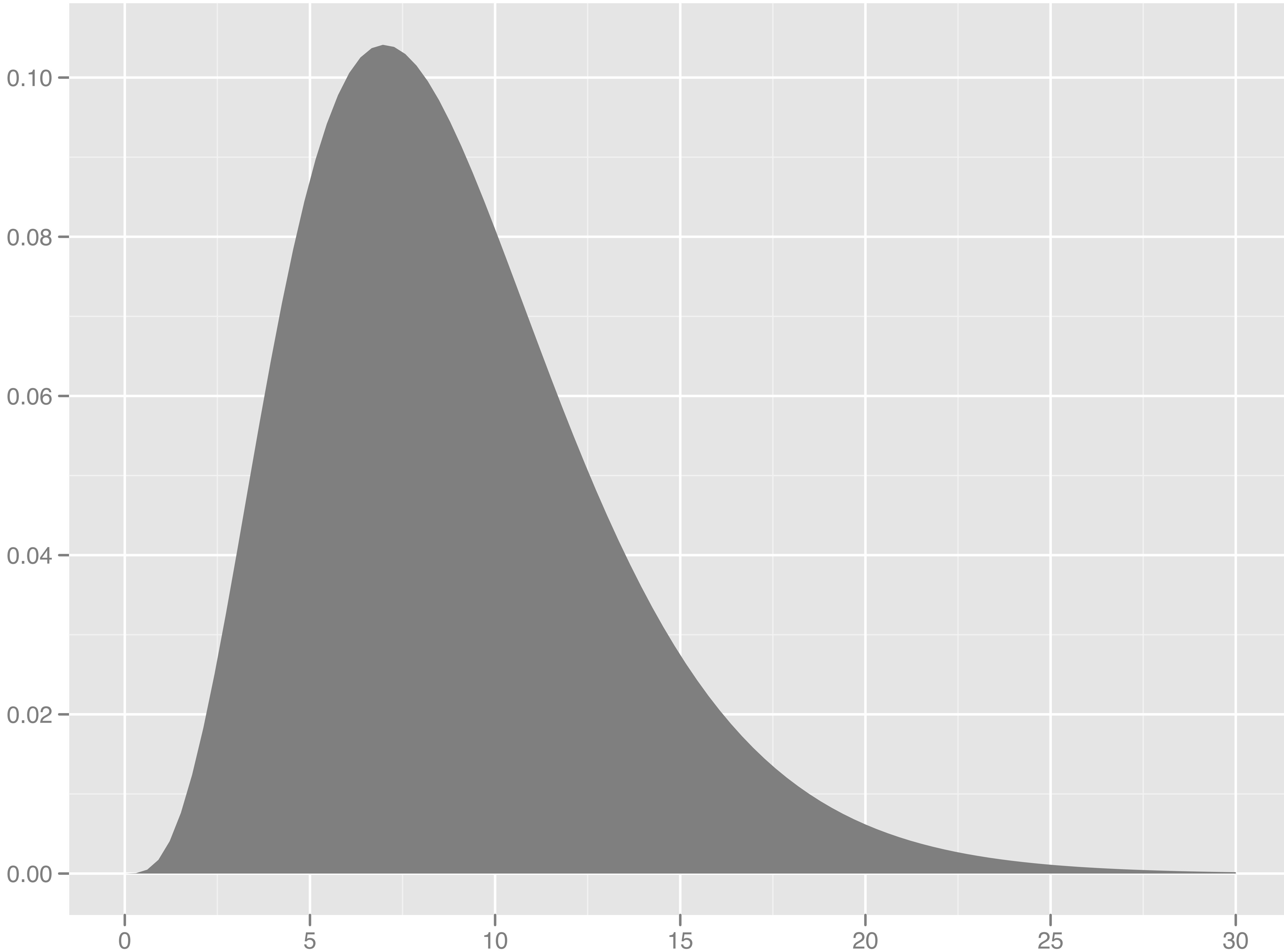
If Q is symmetric, $P(-a < Q < a) = 1 - \alpha$. So $a = F^{-1}(\alpha/2)$, and there is no interval smaller.

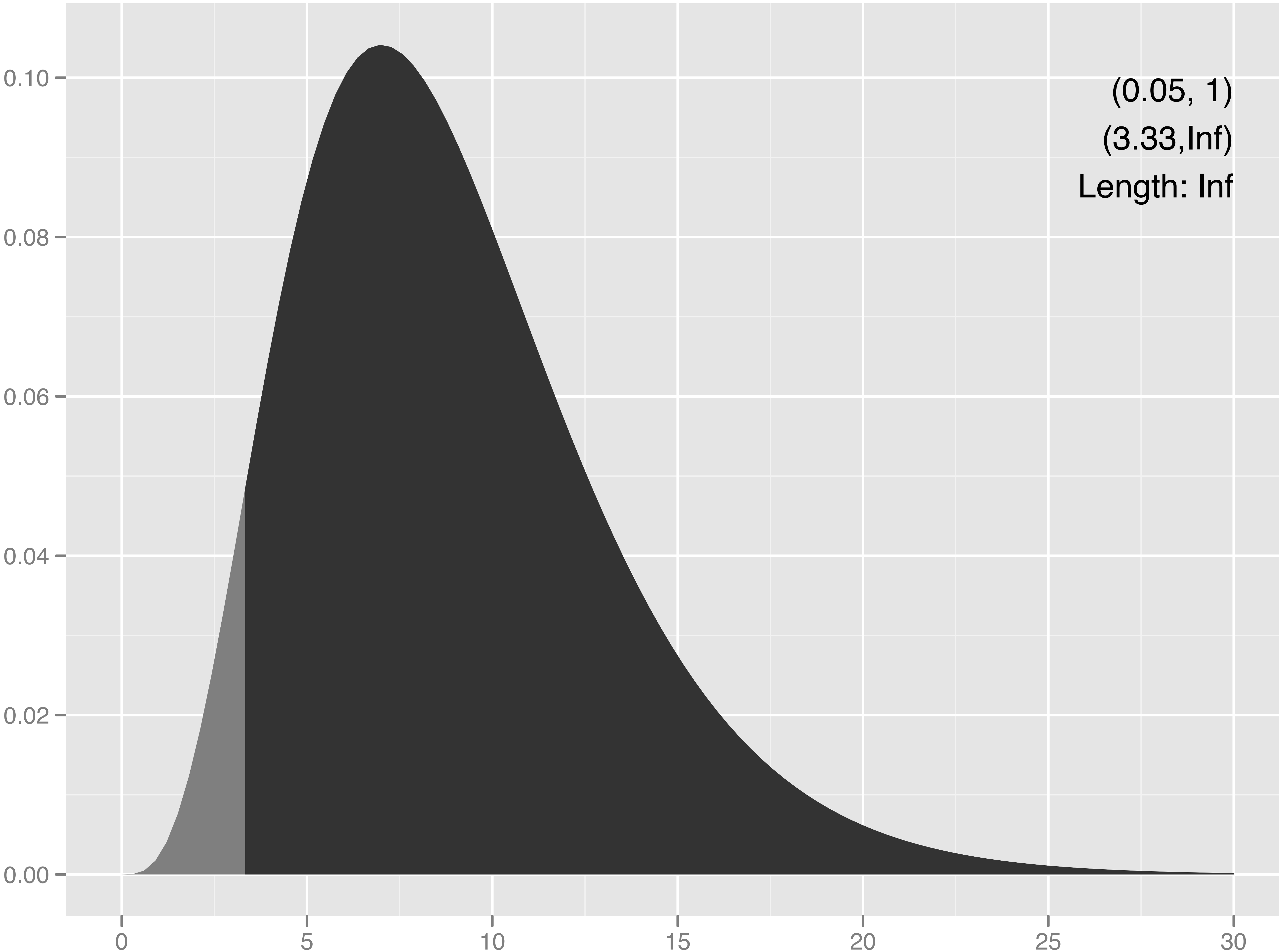
If Q isn't symmetric, pick $a = F^{-1}(\alpha/2)$, $b = F^{-1}(1 - \alpha/2)$, but there might be a shorter interval.

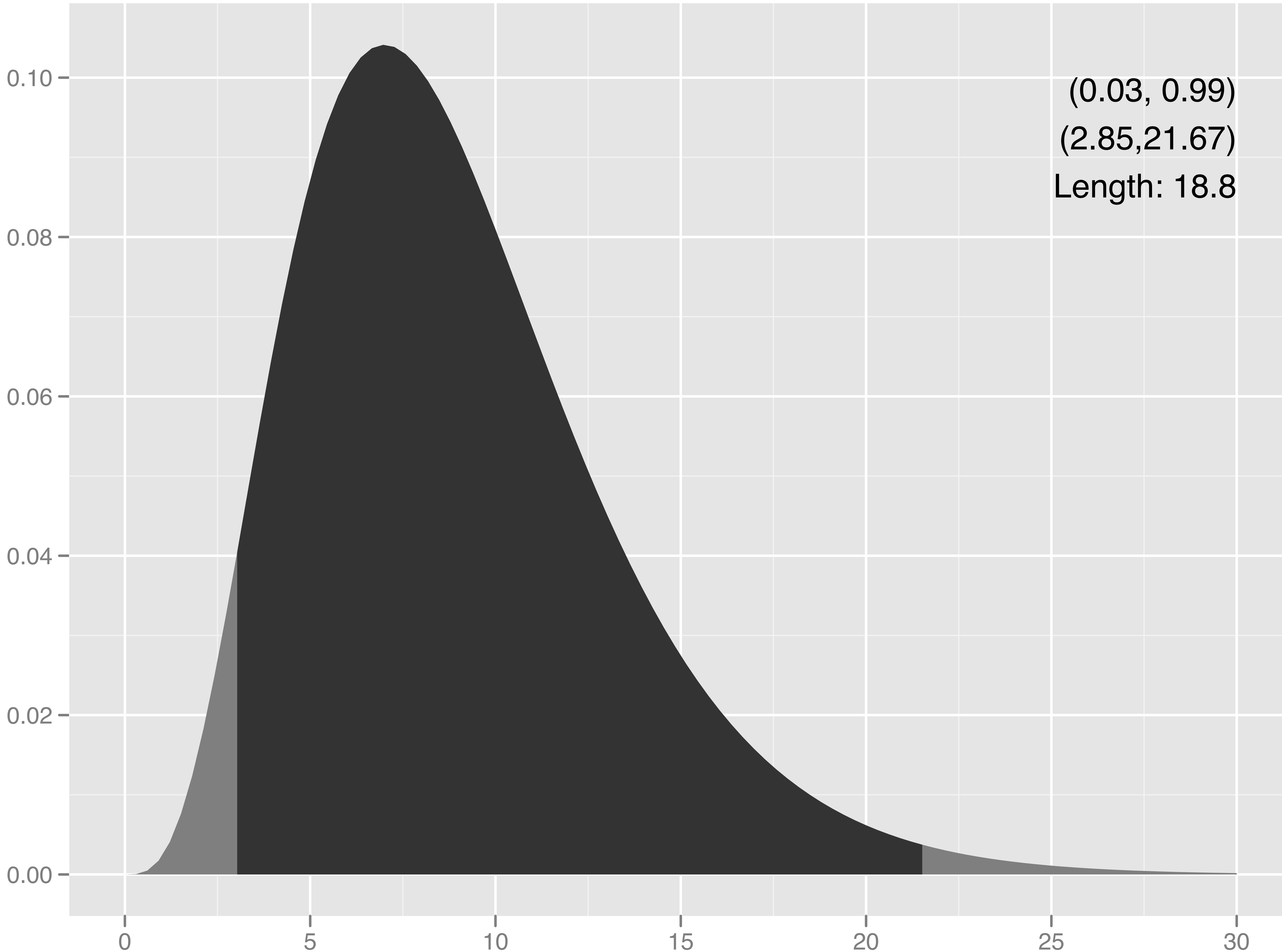
Chi-square

Find confidence interval for $\chi^2(n-1)$. Generally want the shortest confidence interval, but hard to find when not symmetric.

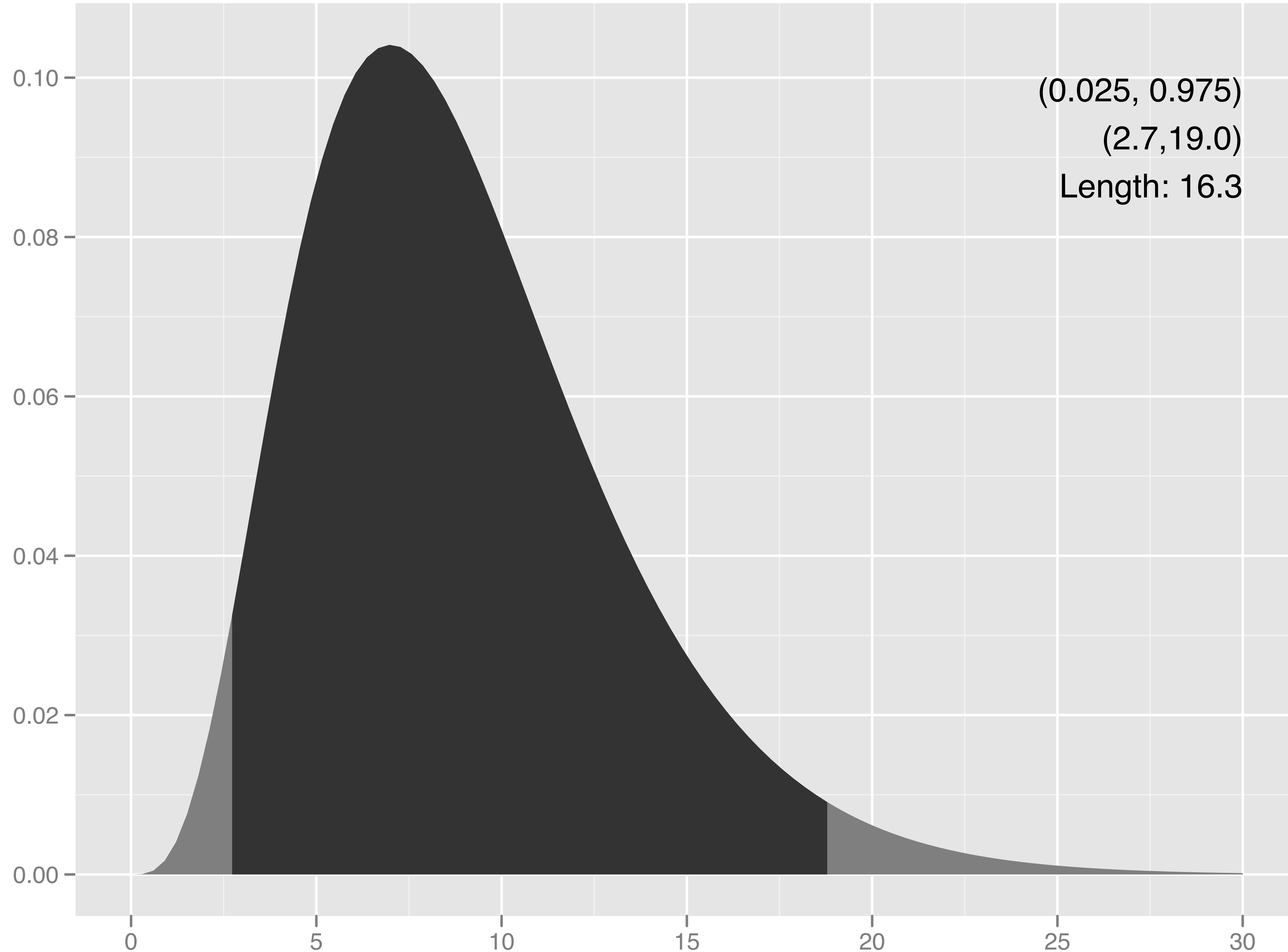
Any of the following are correct. The best has the smallest interval.


0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30

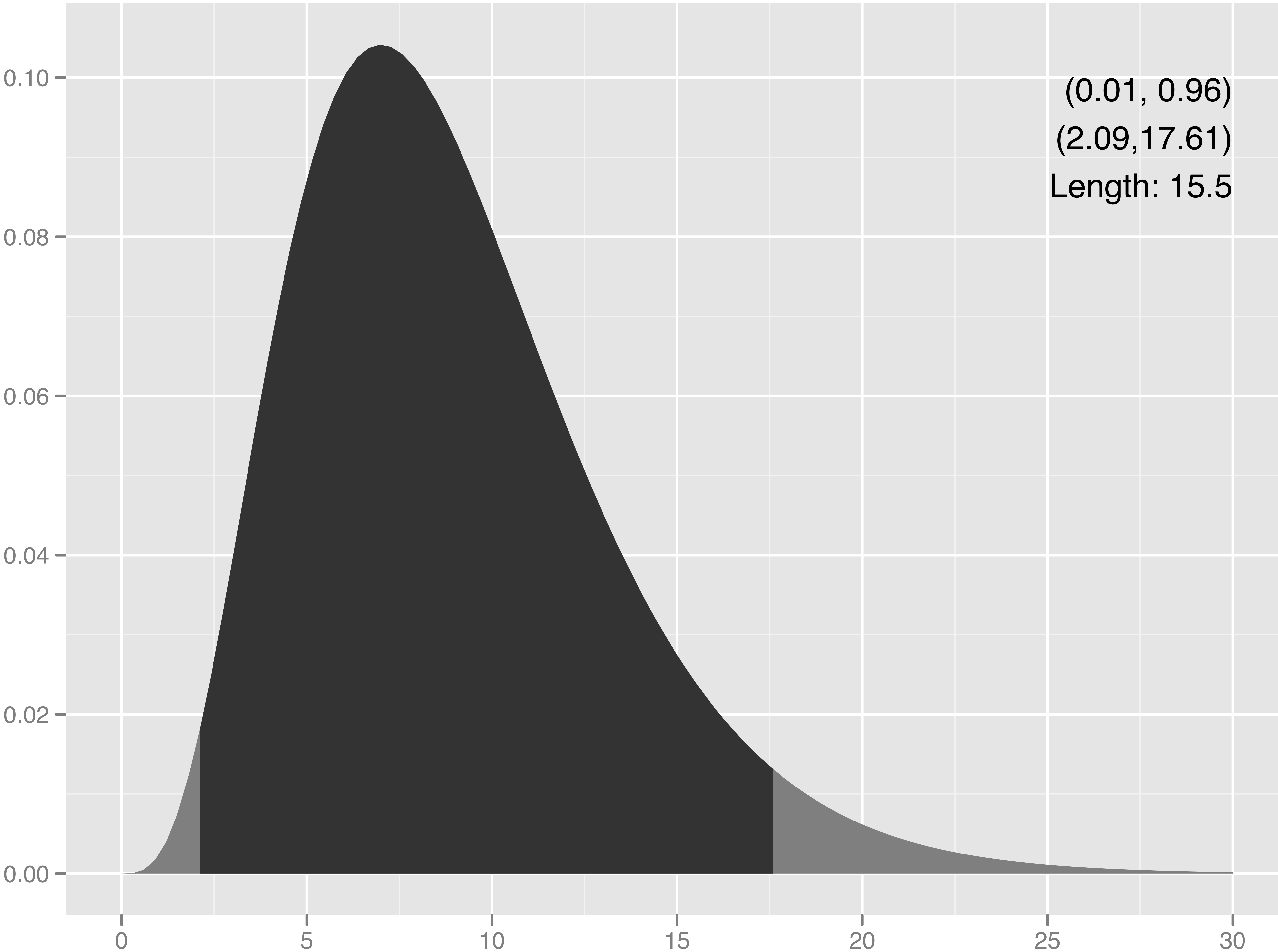

0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30
(0.05, 1)
(3.33,Inf)
Length: Inf

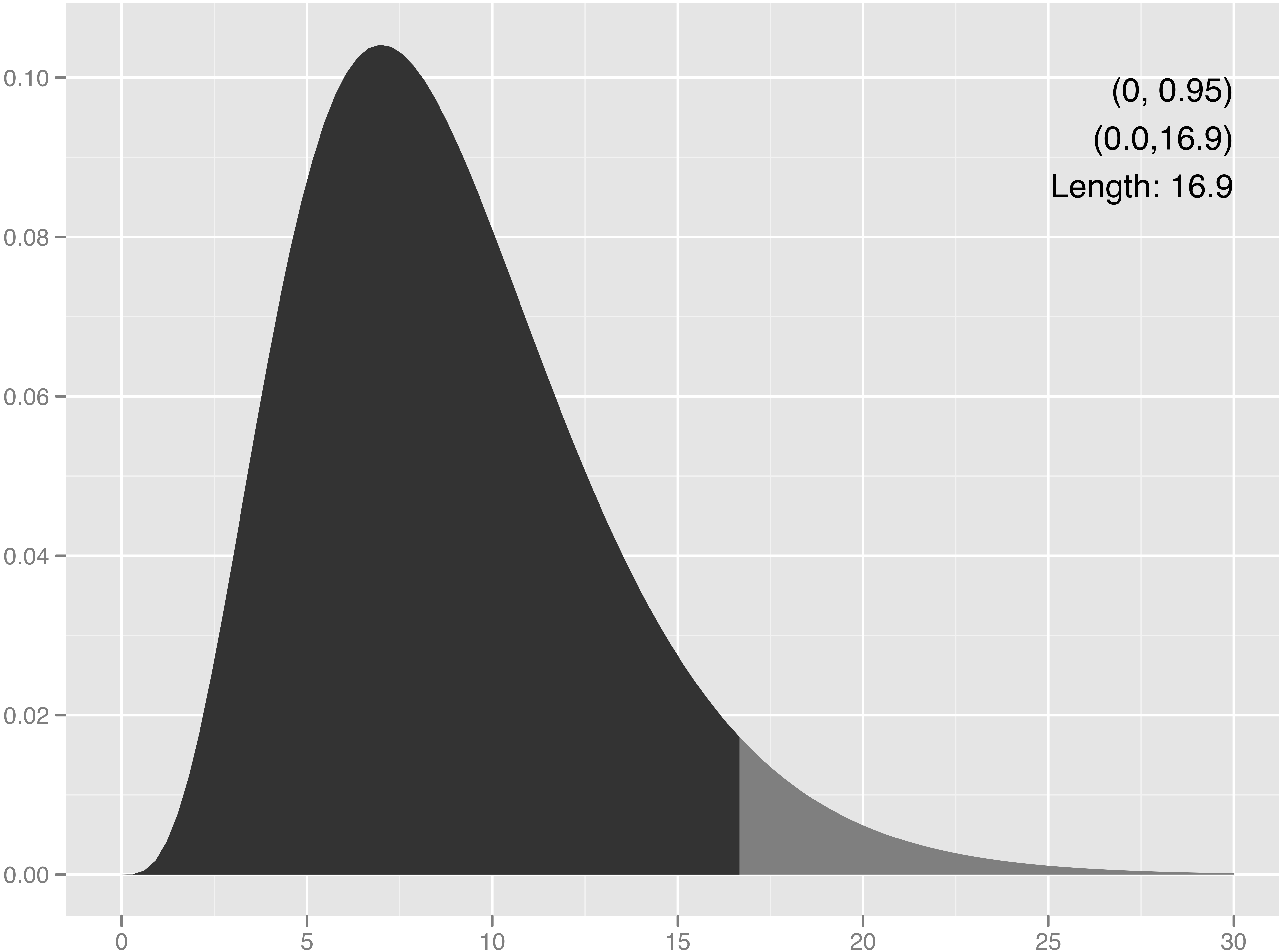


(0.03, 0.99)
(2.85,21.67)
Length: 18.8
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30



(0.025, 0.975)
(2.7,19.0)
Length: 16.3
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30


0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30
(0.01, 0.96)
(2.09,17.61)
Length: 15.5


(0, 0.95)
(0.0,16.9)
Length: 16.9
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30

Example

We want a 90% confidence interval, then two possible ends for the interval are $F^{-1}(0.05)$ and $F^{-1}(0.95)$

Variance

X_i iid Normal(5, σ^2), i = 1, ..., 10

I ran the experiment and recorded the following results: 0.15 1.48 1.25 2.47 1.09 1.95 1.46 1.49 2.81 1.96 (s^2 = 0.55)

Find a 95% confidence interval for the variance. (Hint: If $X \sim \chi^2(9)$, $F_X(19) = 0.975$ and $F_X(2.7) = 0.025$).

EC: Can you also make a confidence interval for the standard deviation?