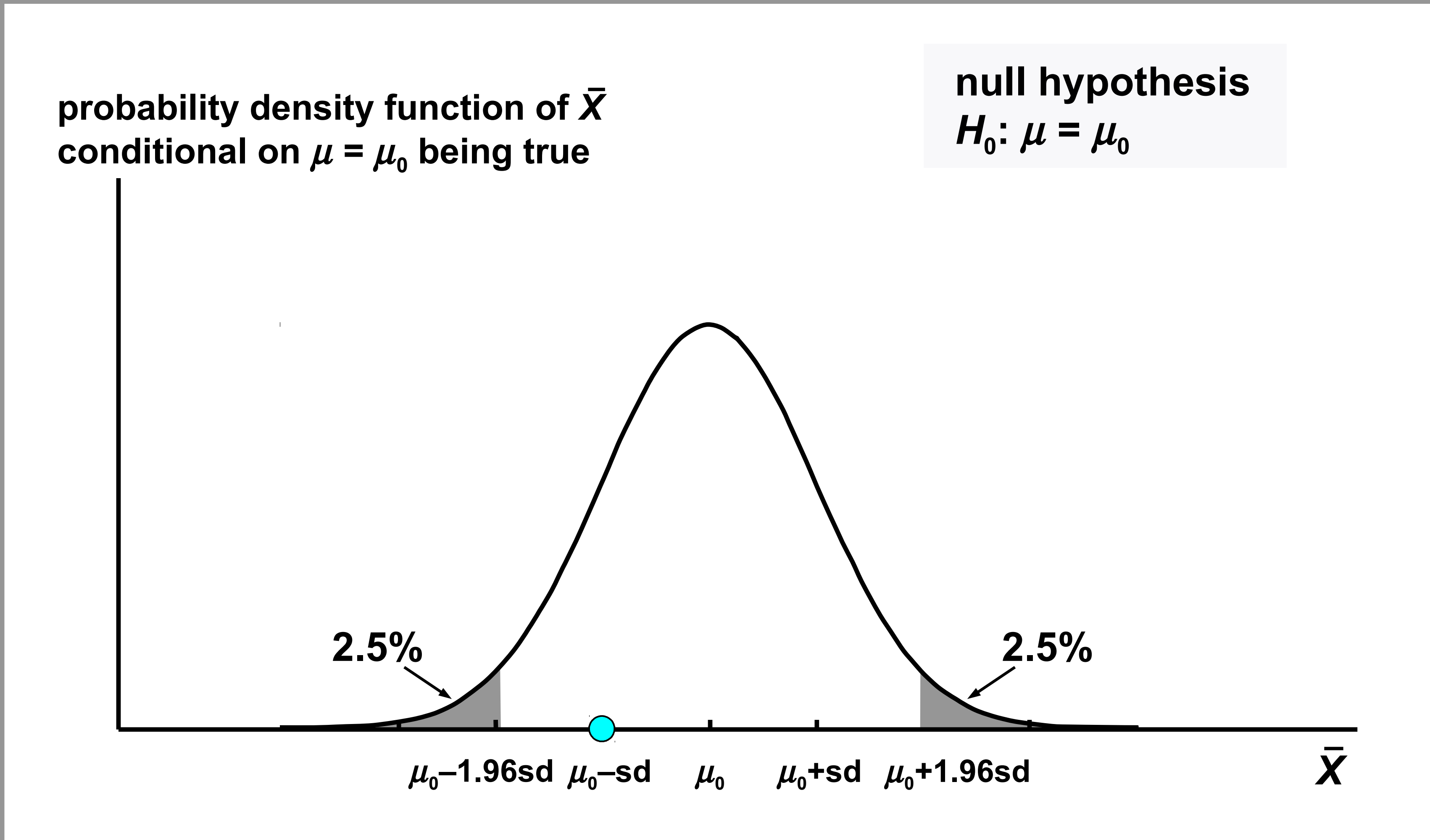
Dougherty

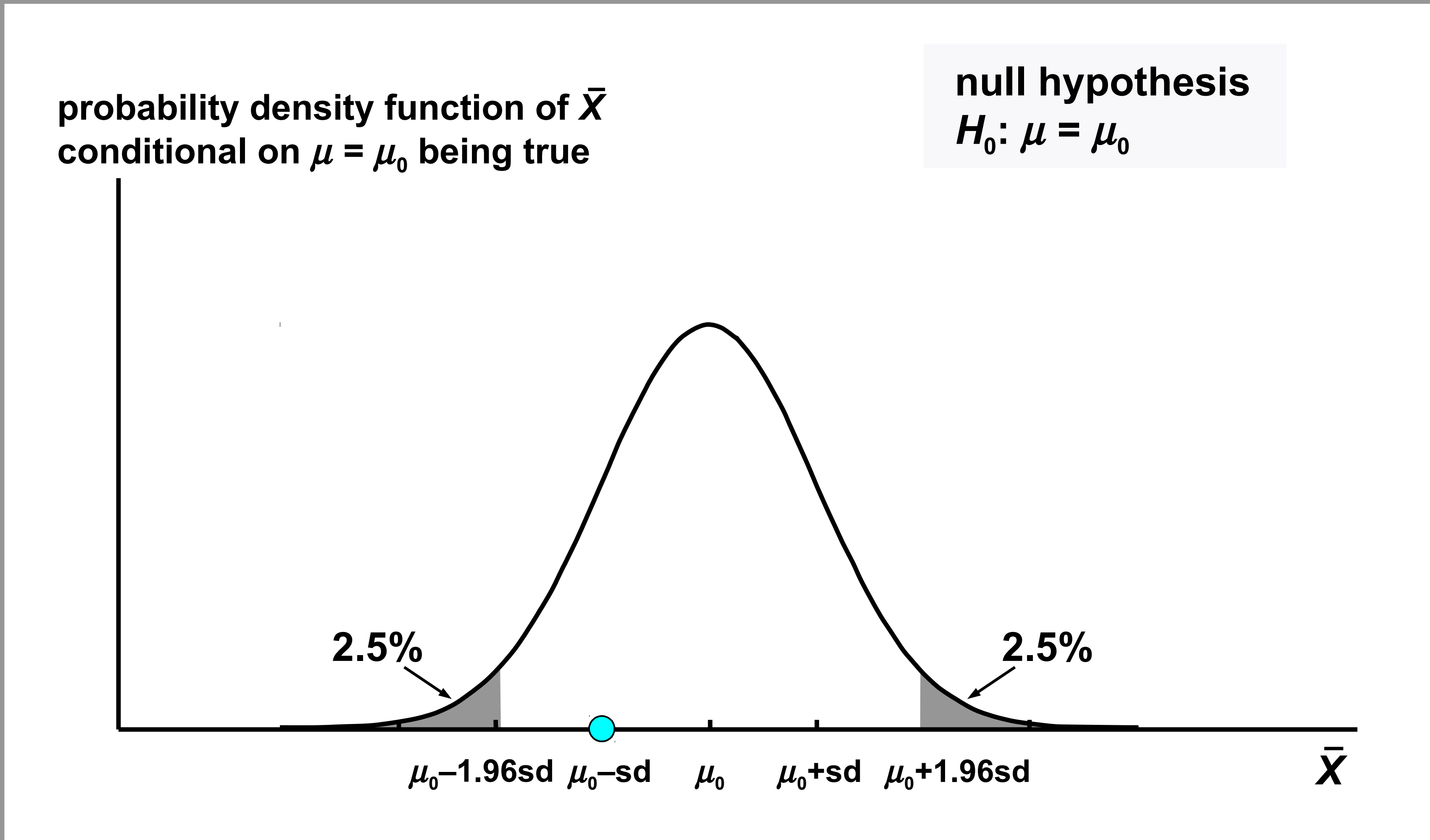
Introduction to Econometrics, 5th edition

Review: Random Variables, Sampling, Estimation, and Inference

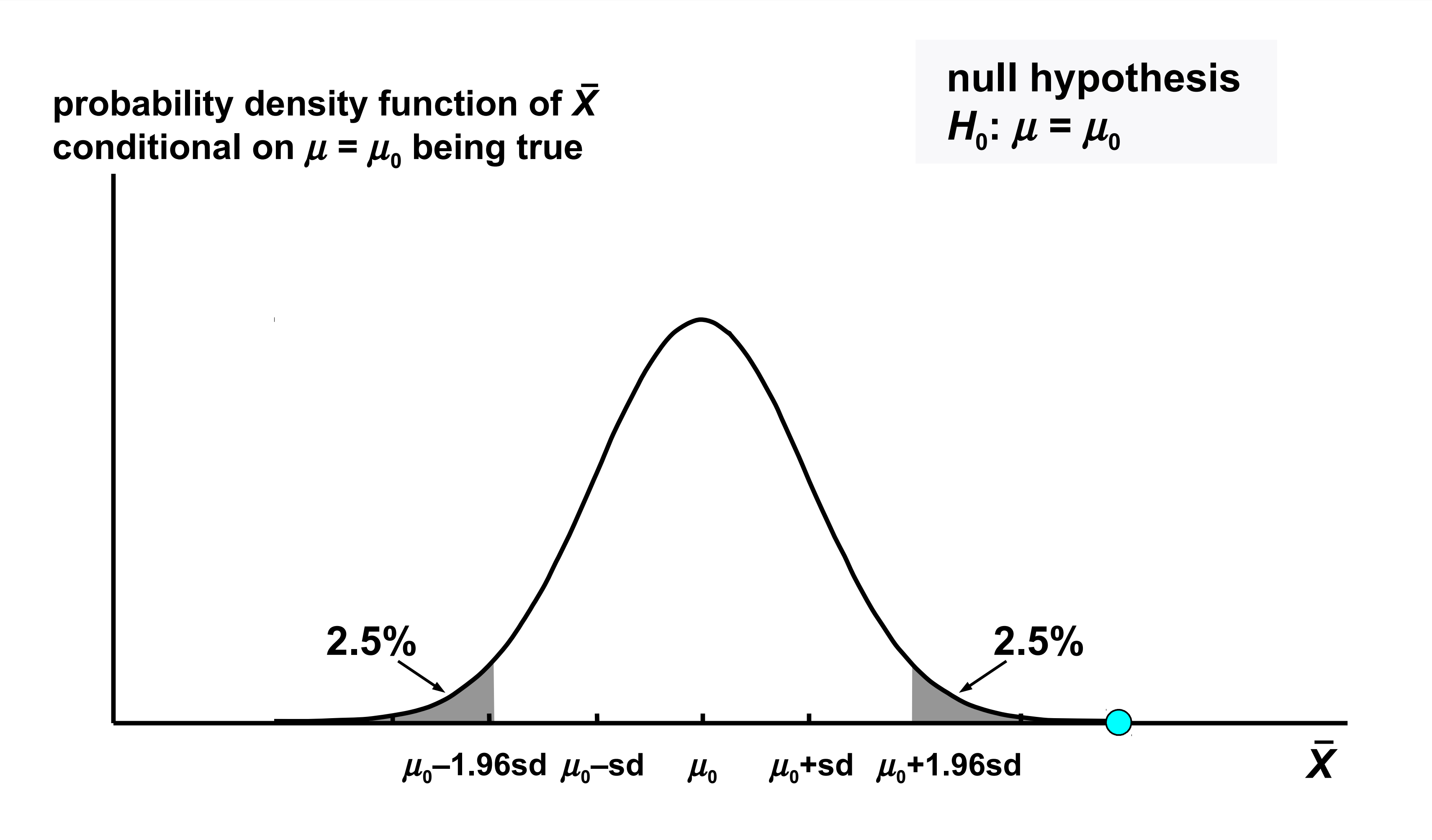
CONFIDENCE INTERVALS



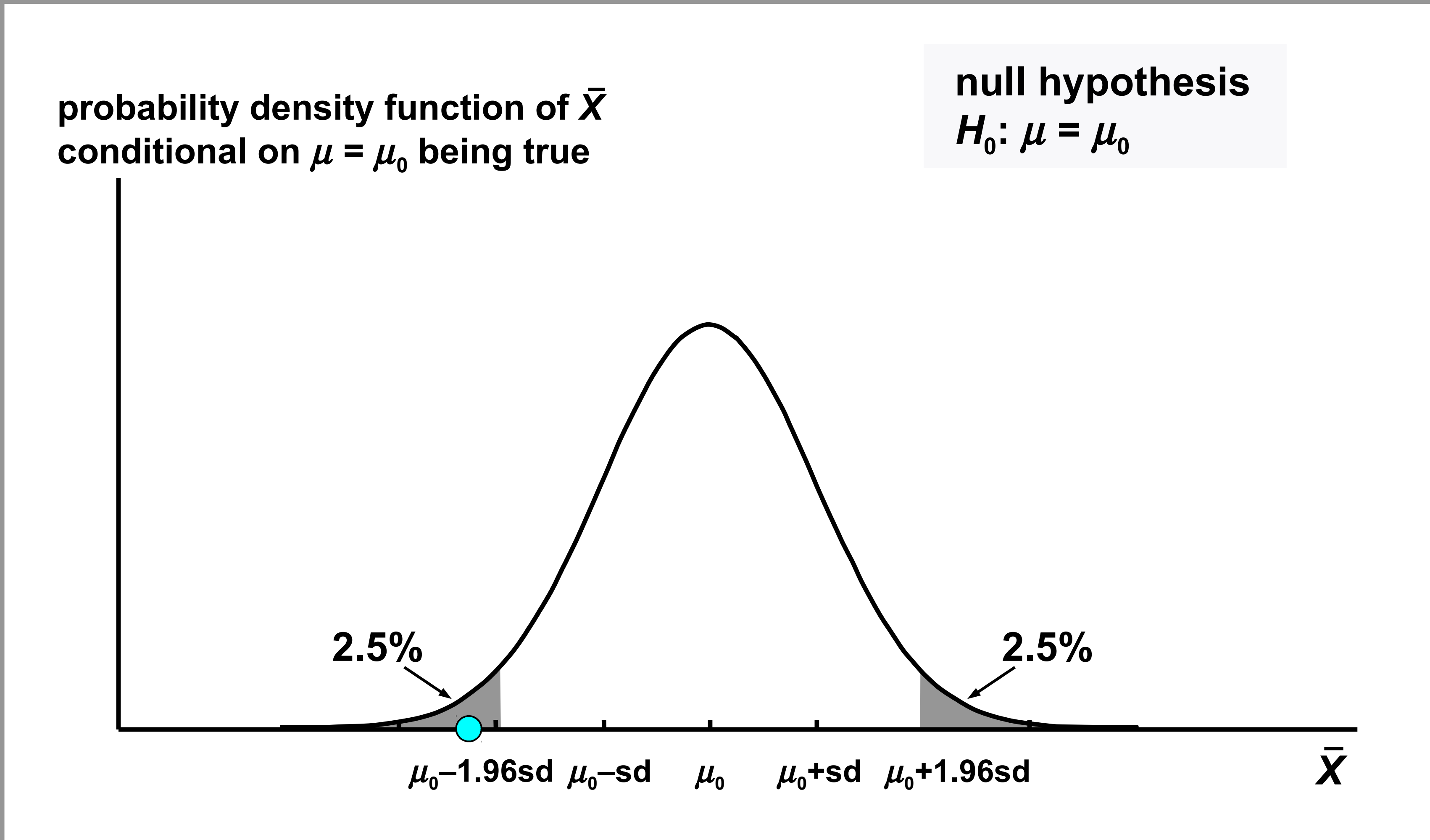
In the sequence on hypothesis testing, we started with a given hypothesis, for example H_0: $\mu = \mu_0$, and considered whether an estimate $\bar{X}$ derived from a sample would or would not lead to its rejection.



Using a 5% significance test, this estimate would not lead to the rejection of H_0.

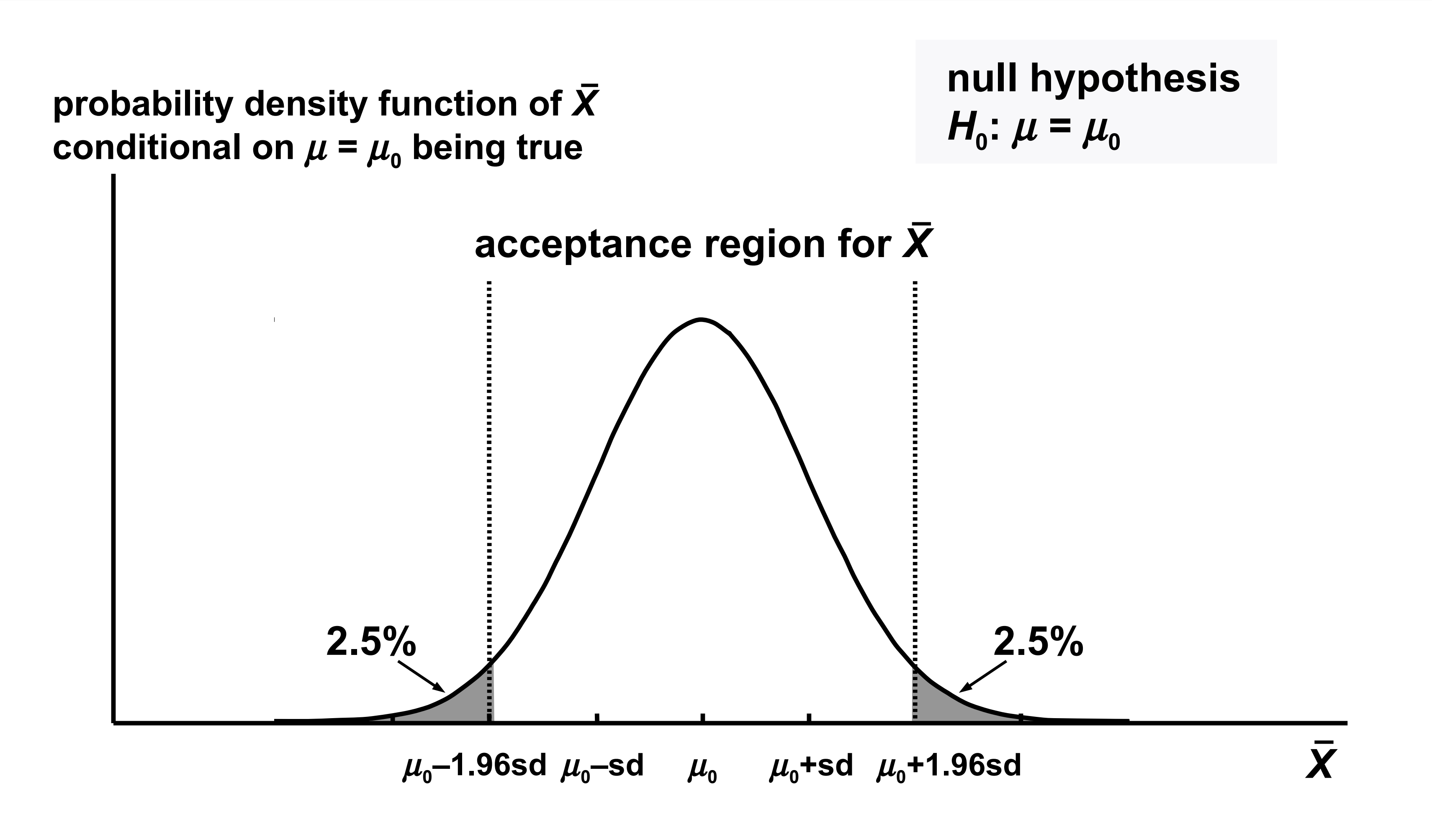


This estimate would cause H_0 to be rejected.



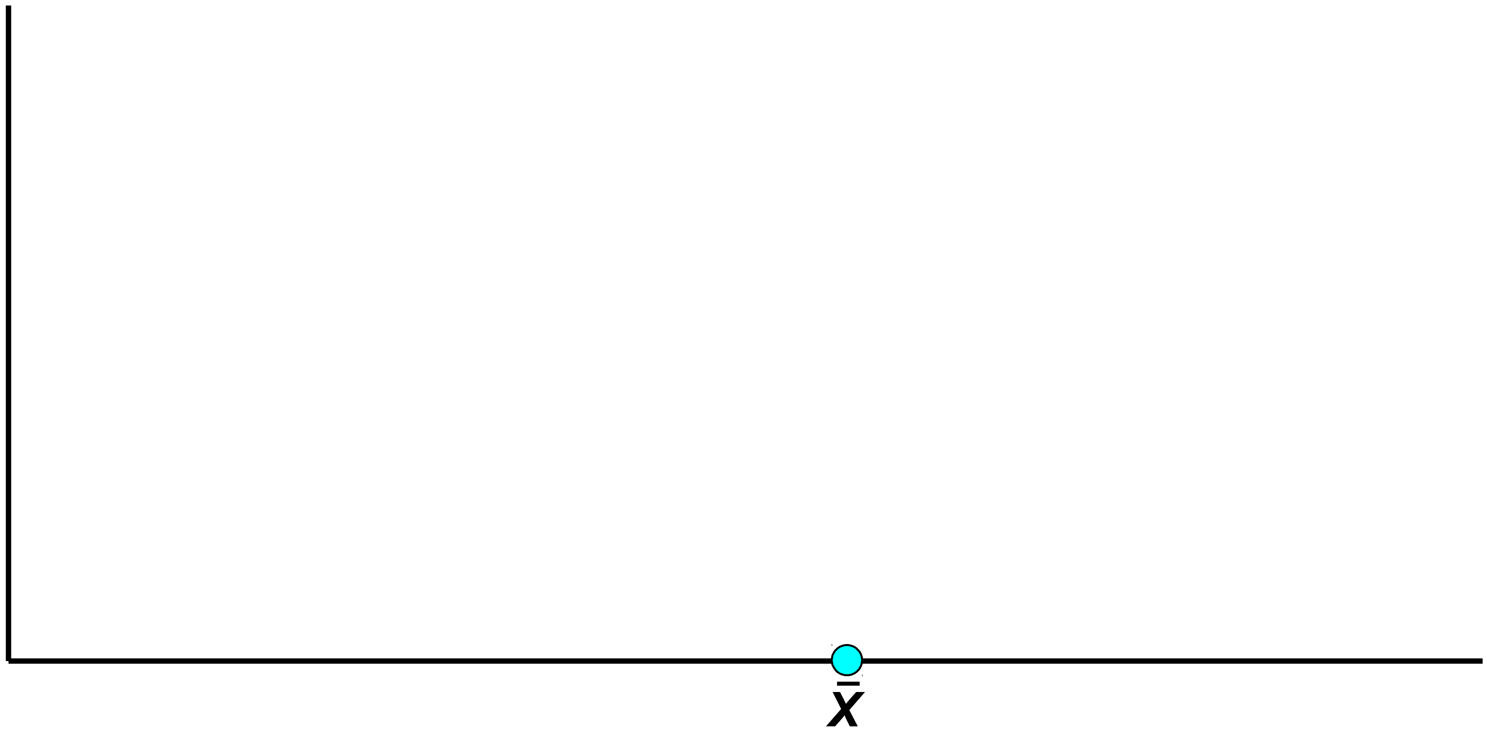
So would this one.

CONFIDENCE INTERVALS



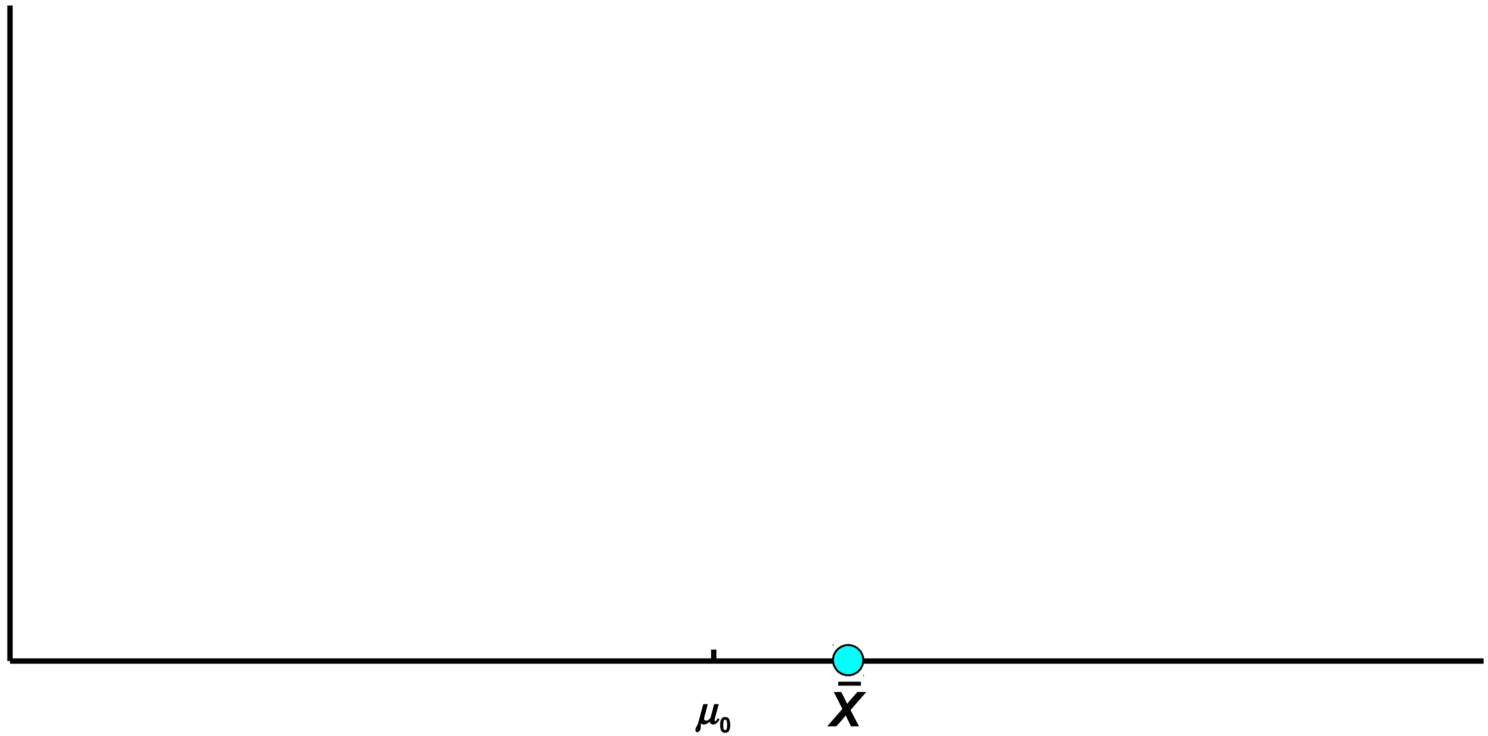
We ended by deriving the range of estimates that are compatible with H_0 and called it the acceptance region.

CONFIDENCE INTERVALS



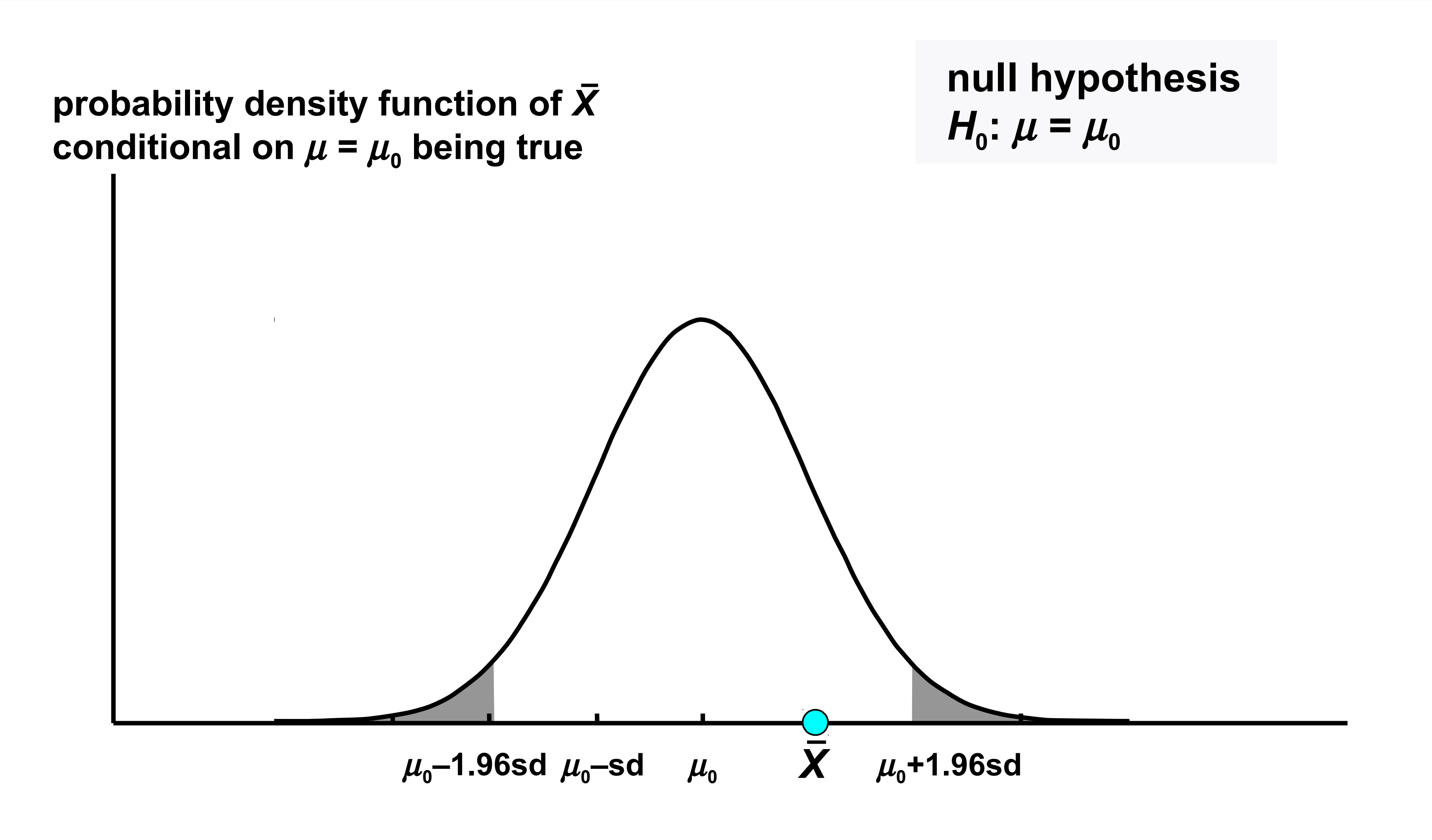
Now we will do exactly the opposite. Given a sample estimate, we will find the set of hypotheses that would not be contradicted by it, using a two-tailed 5% significance test.

null hypothesis
$H_0: \mu = \mu_0$



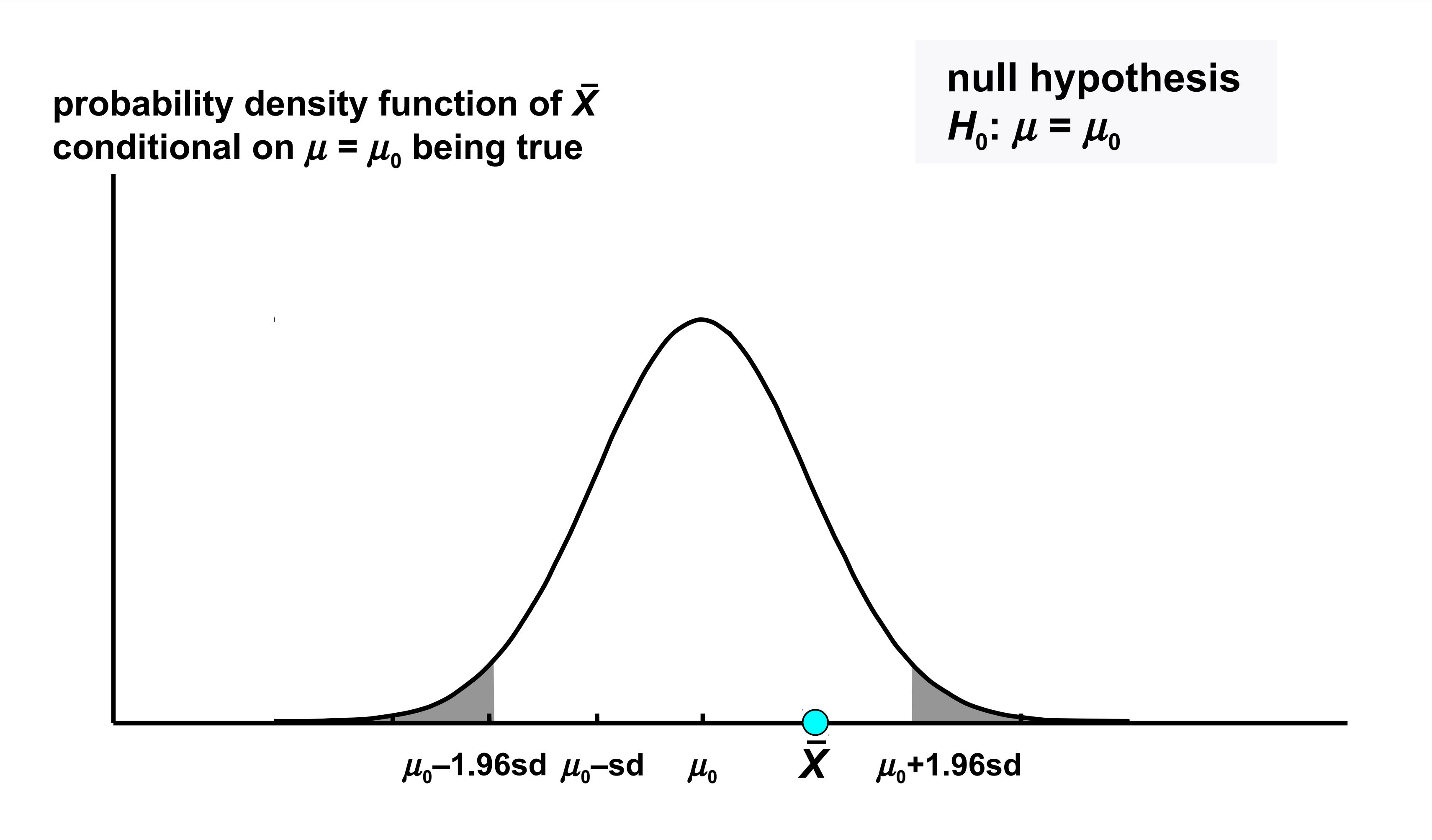
Suppose that someone came up with the hypothesis $H_0: \mu = \mu_0$.

CONFIDENCE INTERVALS



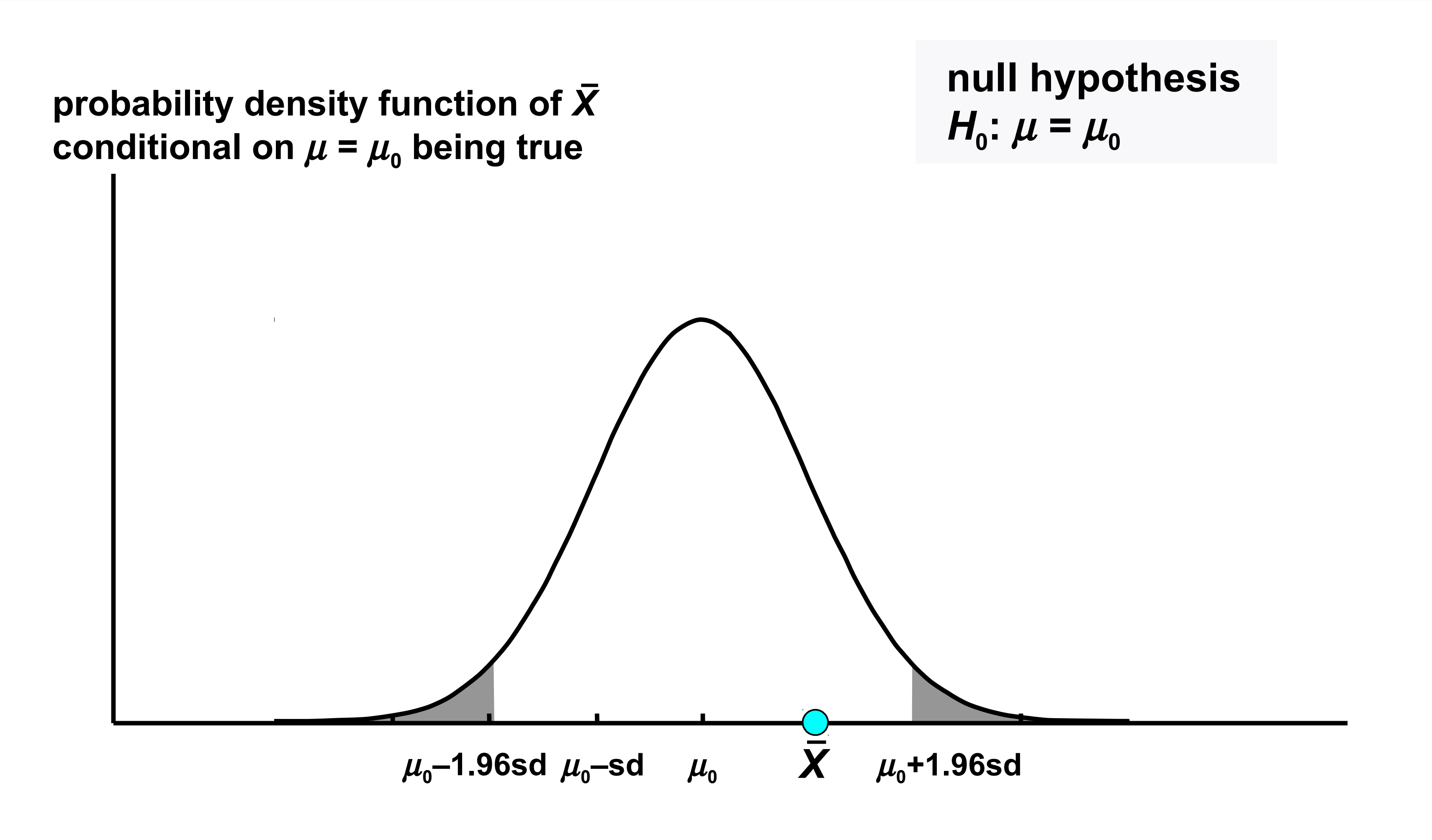
To see if it is compatible with our sample estimate $\bar{X}$, we need to draw the probability distribution of $\bar{X}$ conditional on H_0: $\mu = \mu_0$ being true.

CONFIDENCE INTERVALS



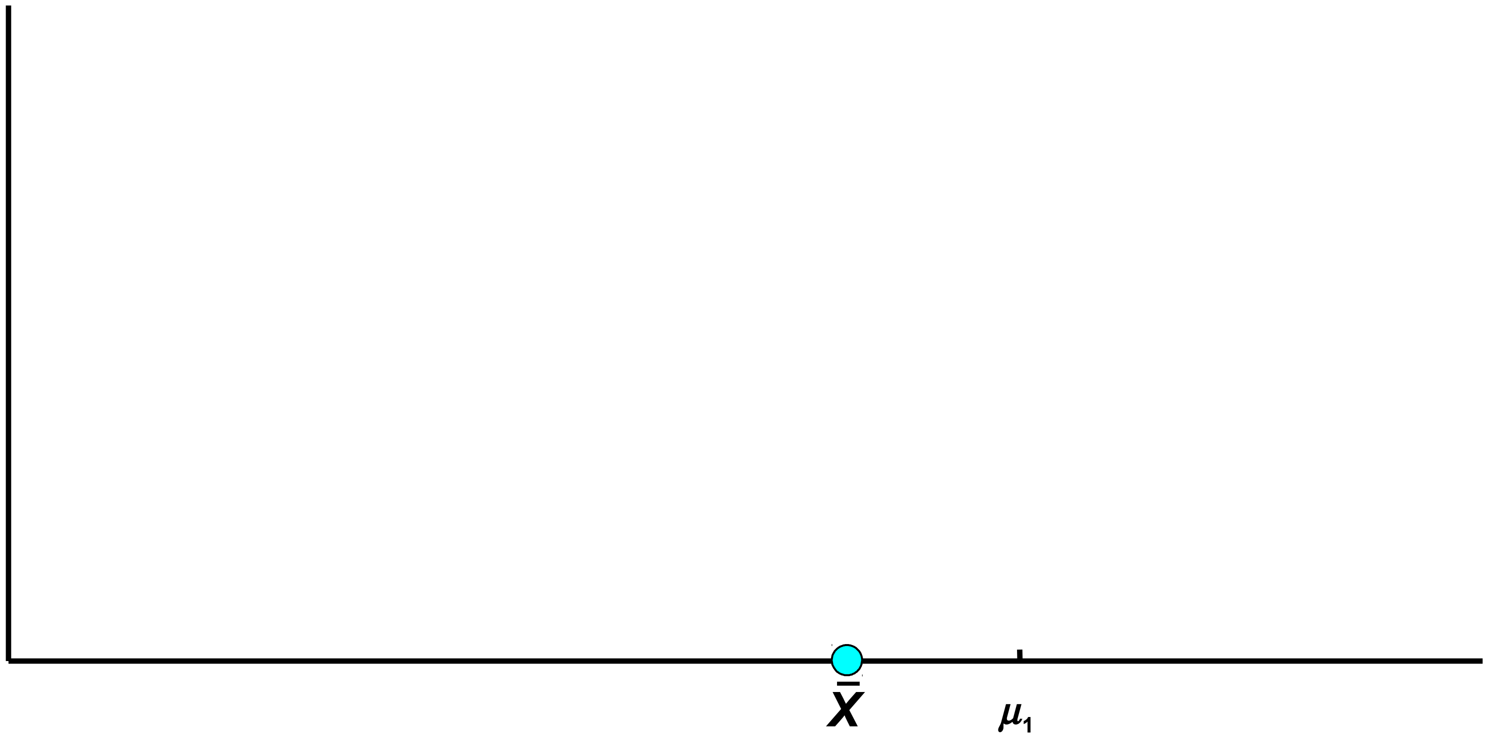
To do this, we need to know the standard deviation of the distribution. For the time being we will assume that we do know it, although in practice we have to estimate it.

CONFIDENCE INTERVALS

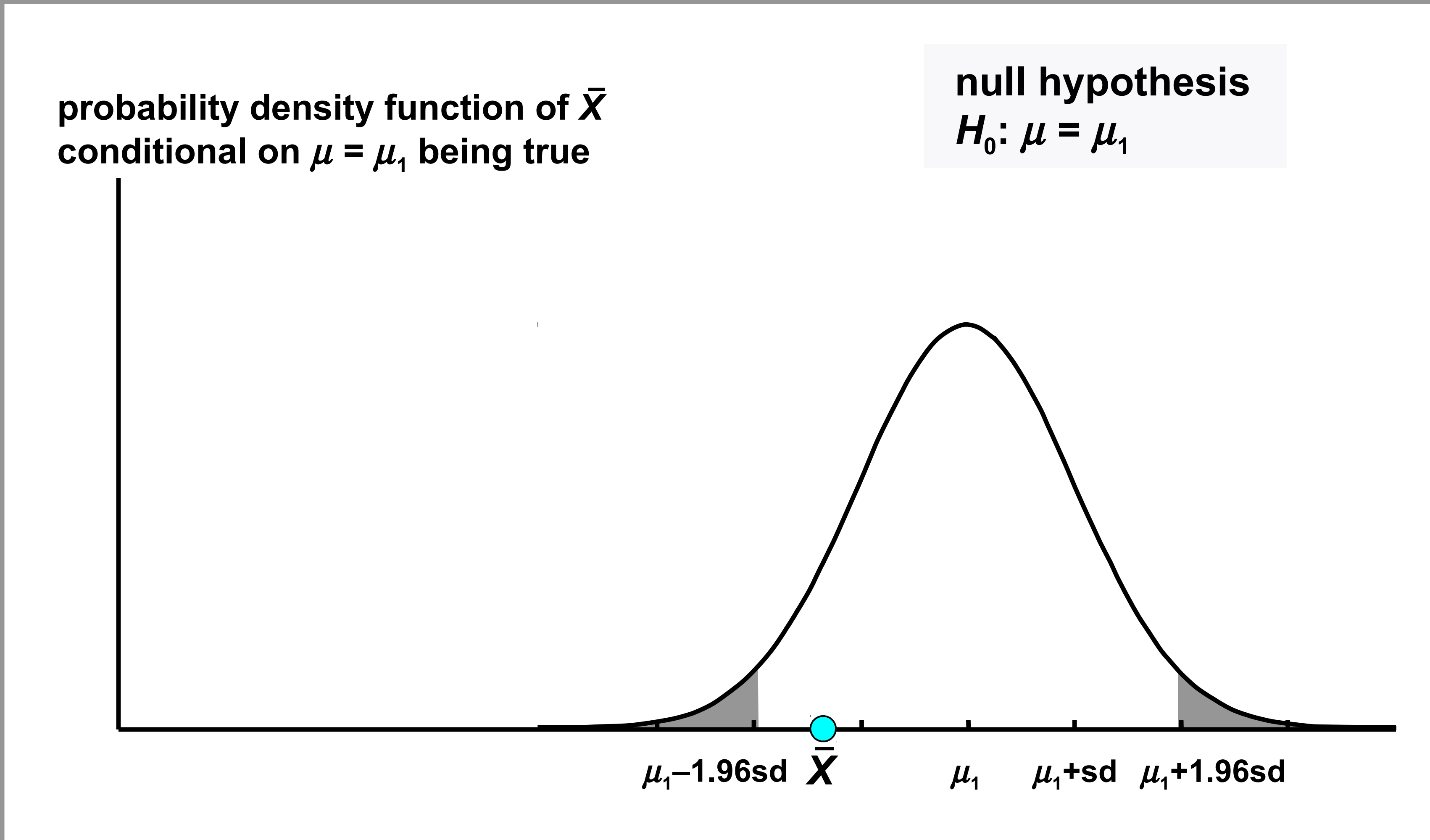


Having drawn the distribution, it is clear this null hypothesis is not contradicted by our sample estimate.

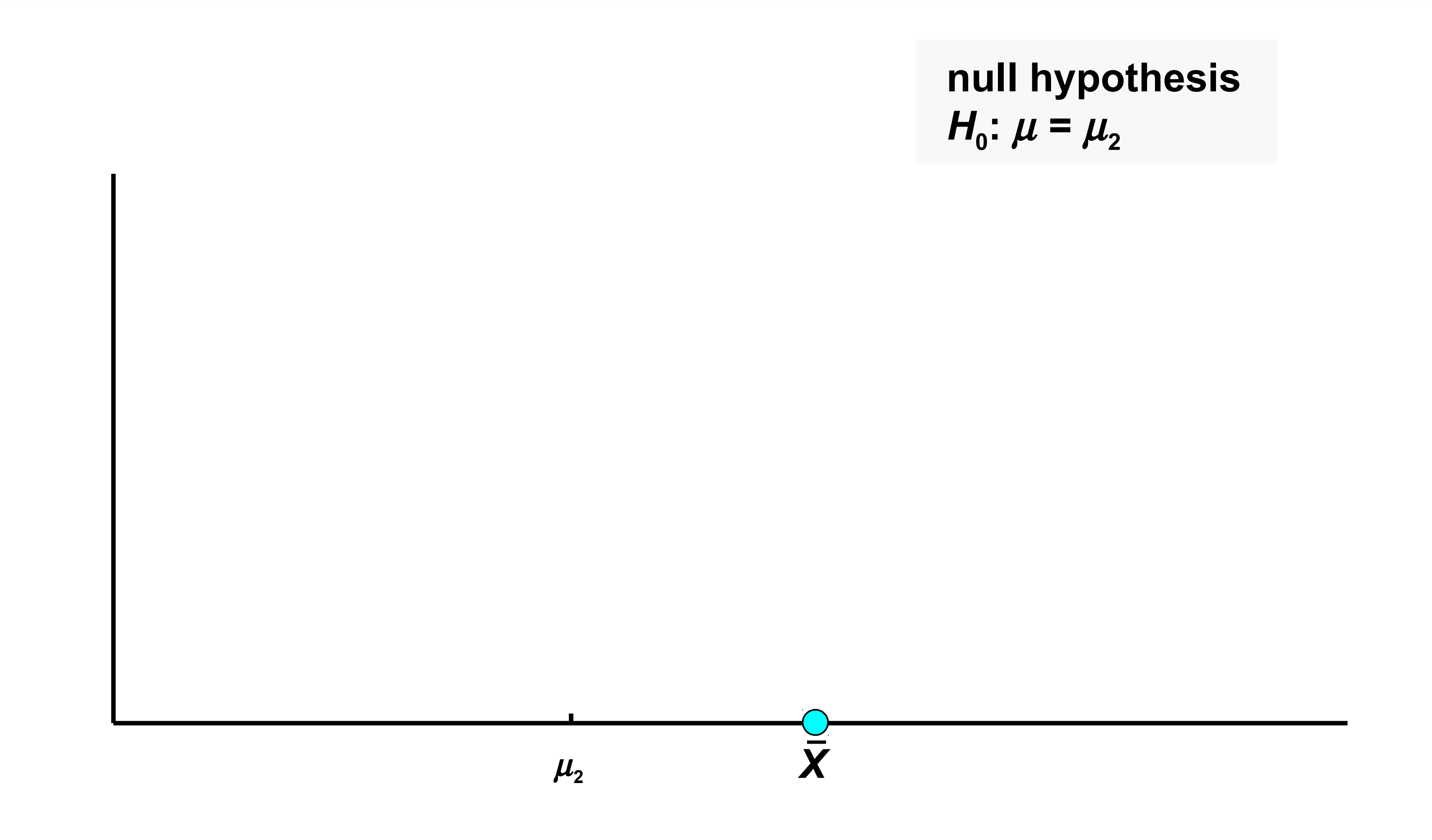
null hypothesis
H_0: $\mu = \mu_1$



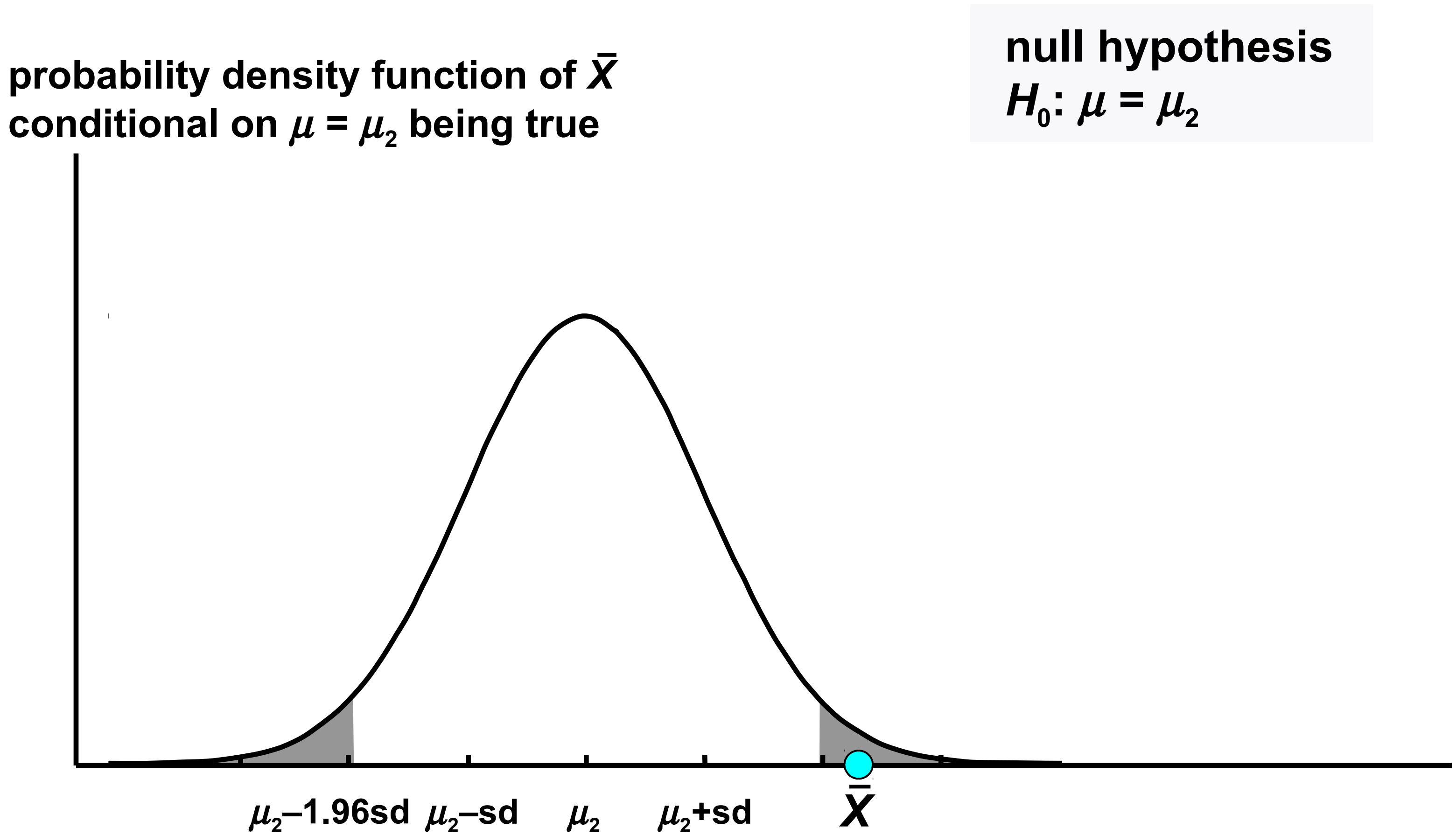
Here is another null hypothesis to consider.



Having drawn the distribution of $\bar{X}$ conditional on this hypothesis being true, we can see that this hypothesis is also compatible with our estimate.

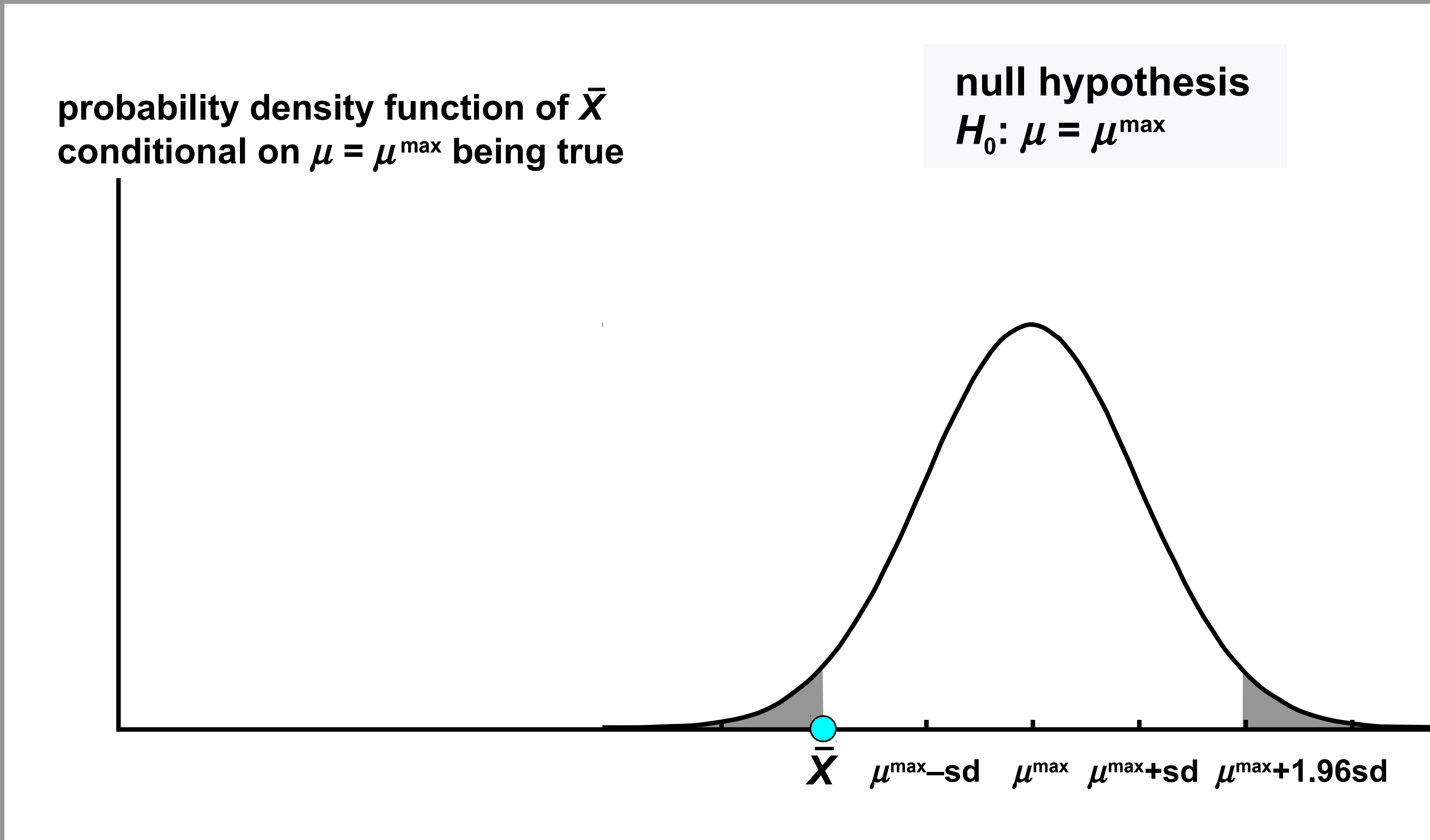


Here is another null hypothesis.



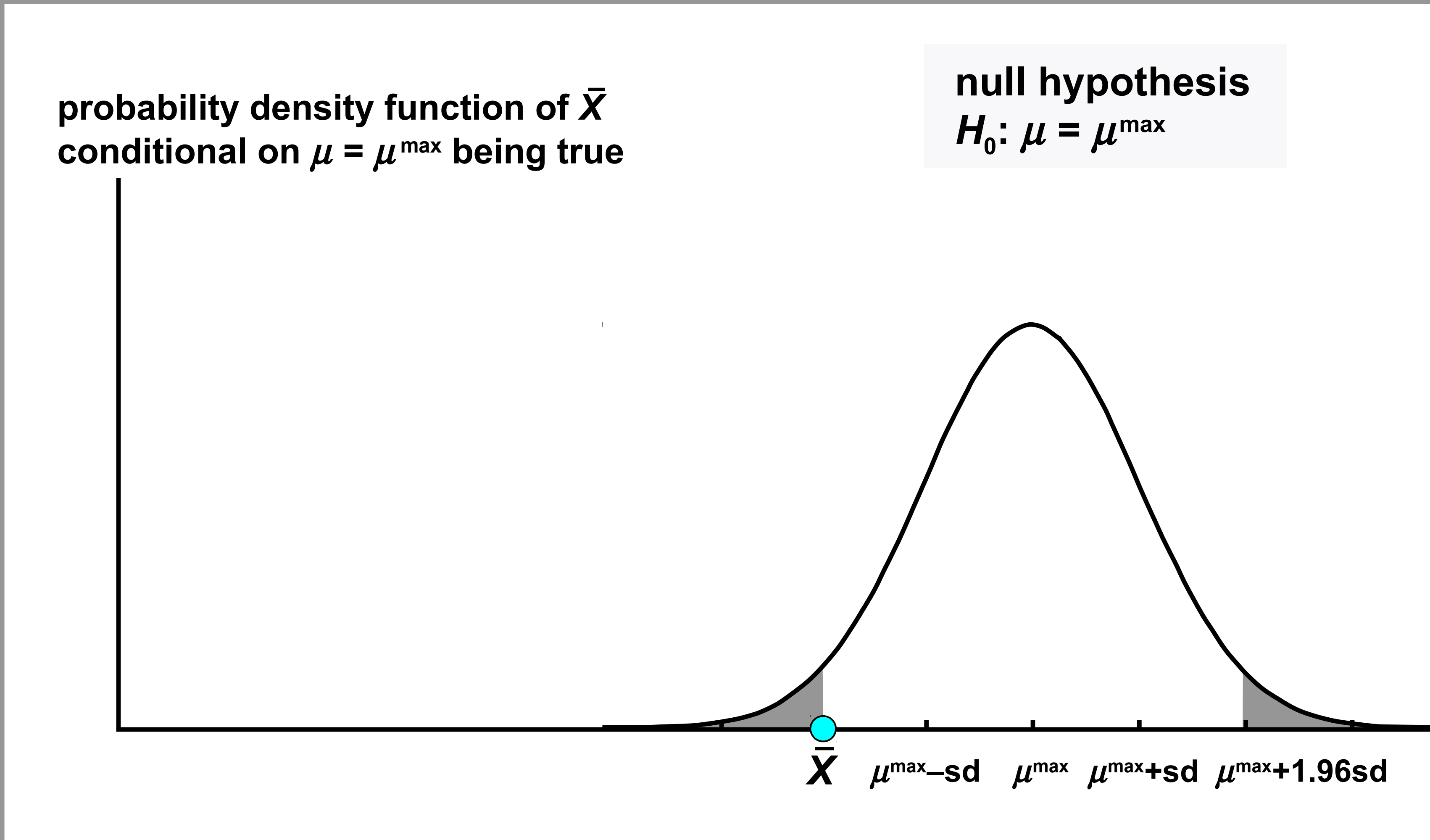
It is incompatible with our estimate because the estimate would lead to its rejection.

CONFIDENCE INTERVALS



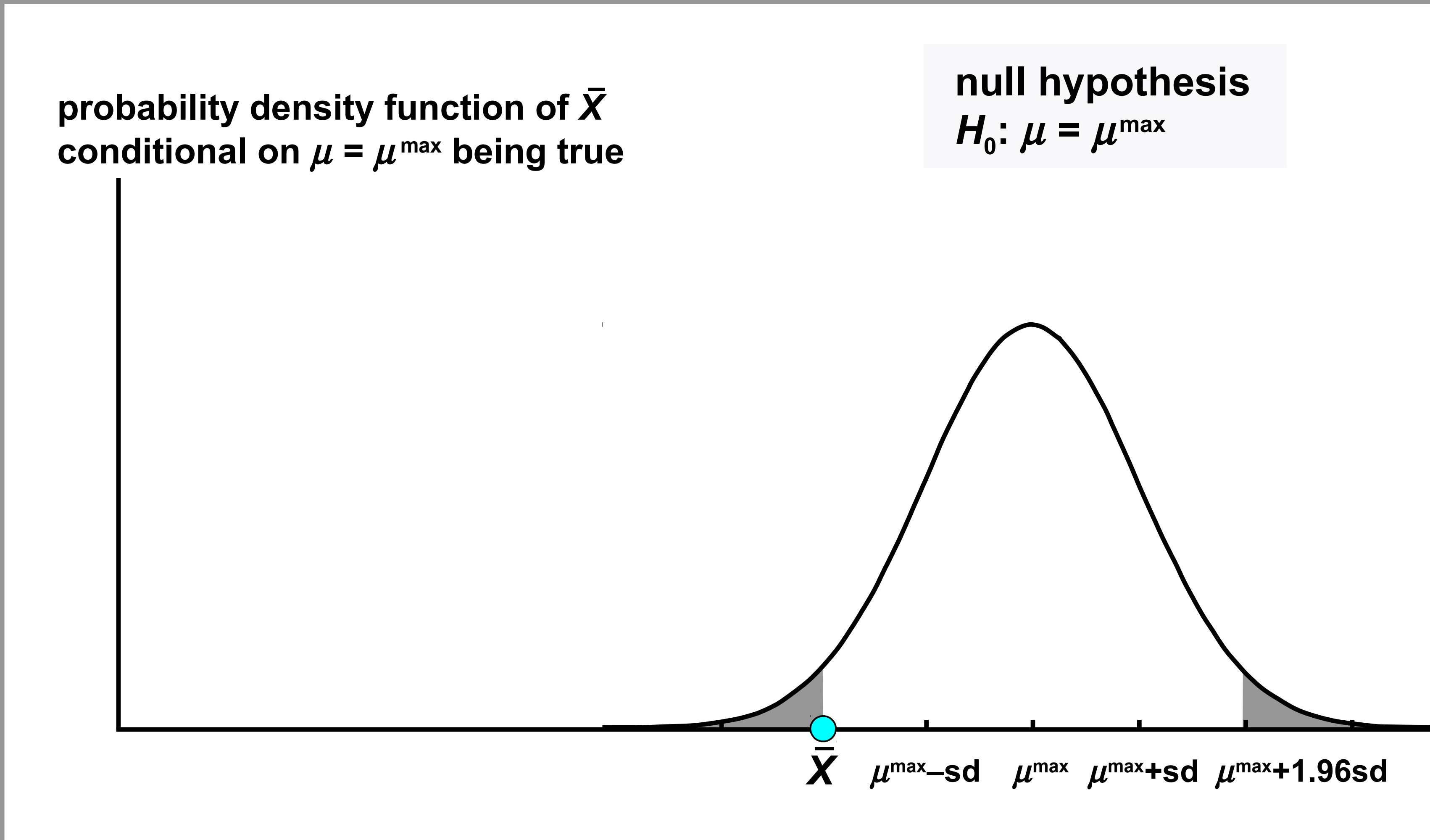
The highest hypothetical value of μ not contradicted by our sample estimate is shown in the diagram.

CONFIDENCE INTERVALS

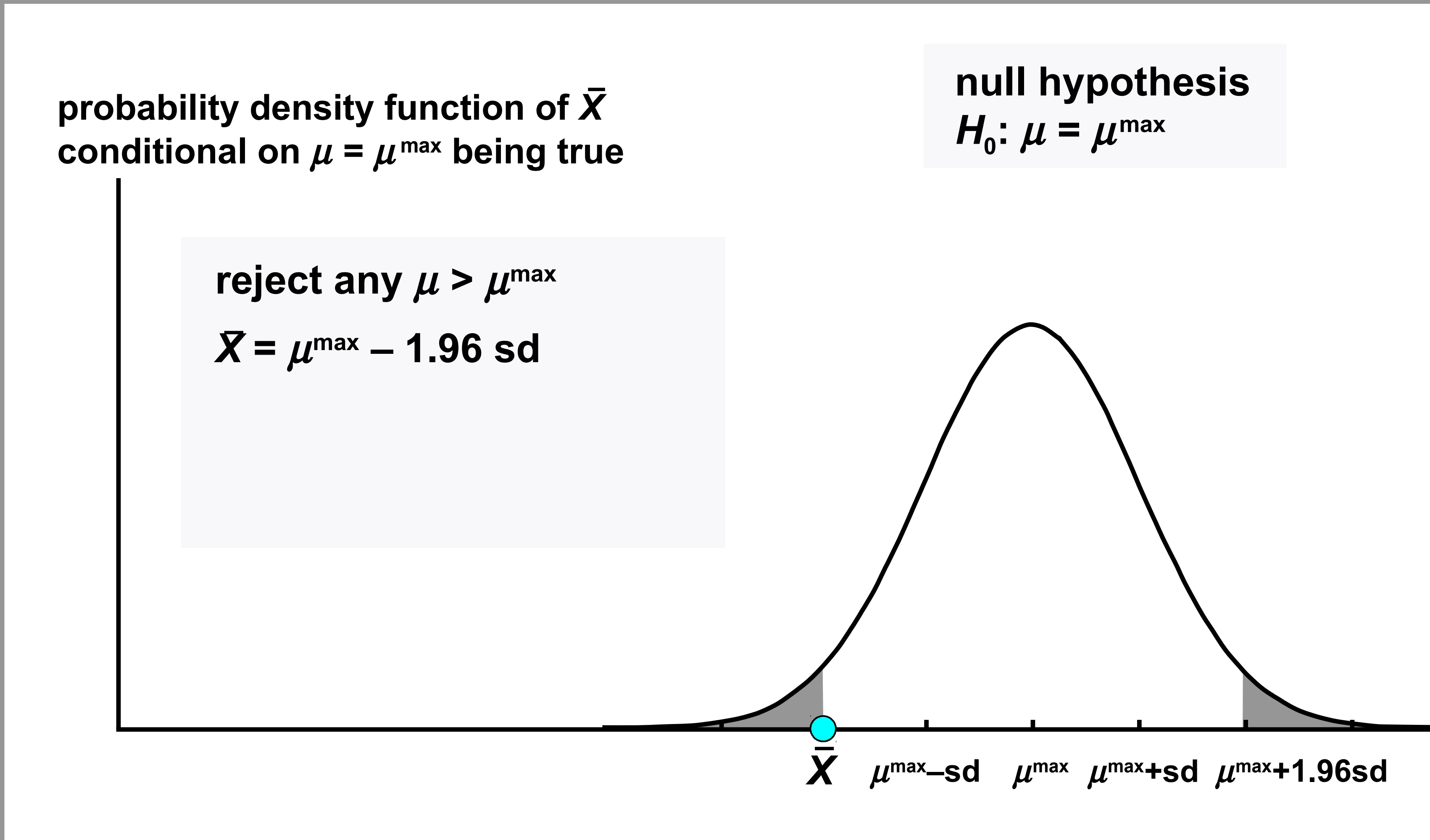


When the probability distribution of $\bar{X}$ conditional on it is drawn, it implies $\bar{X}$ lies on the edge of the left 2.5% tail. We will call this maximum value μ^{max}.

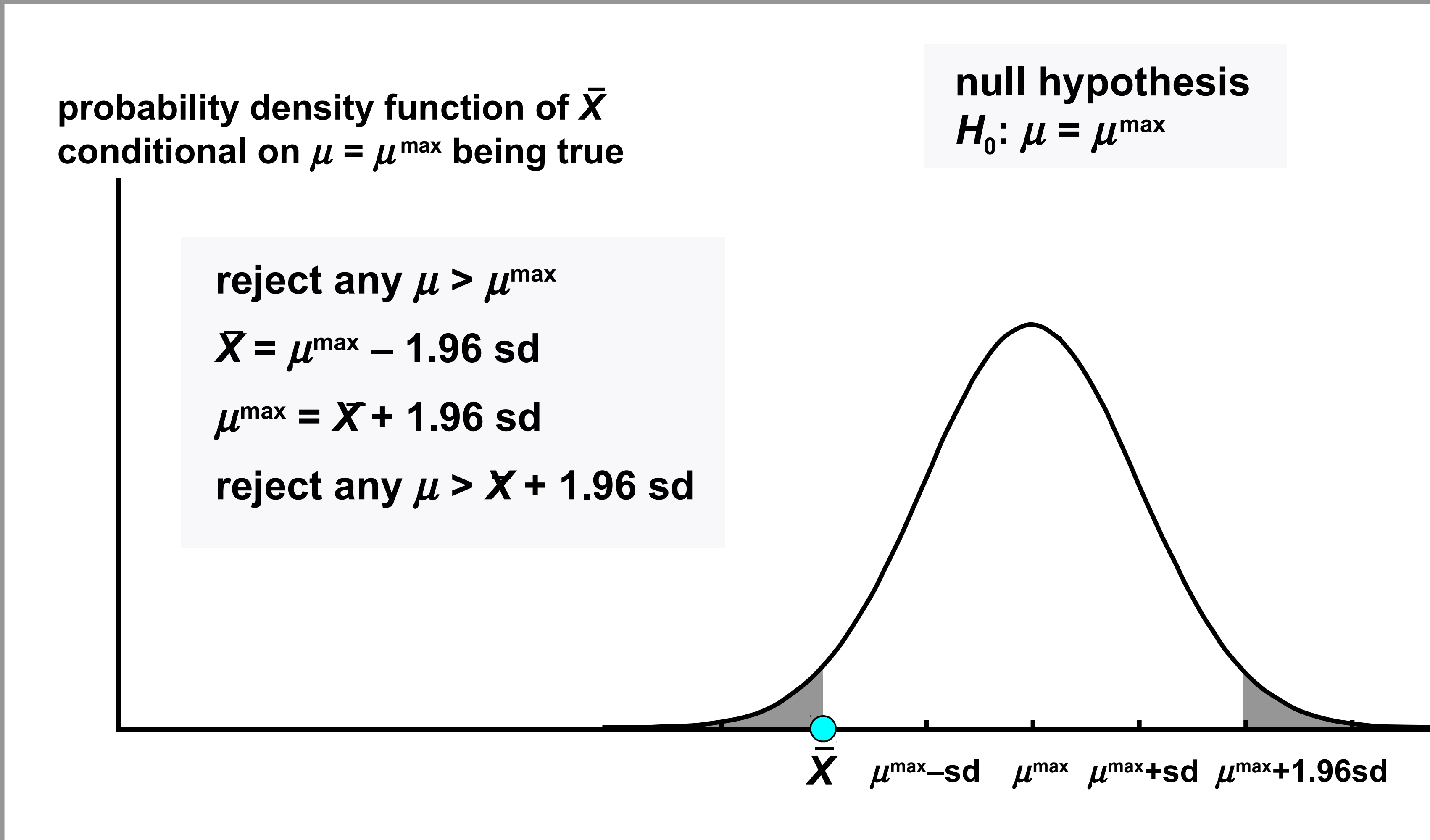
CONFIDENCE INTERVALS



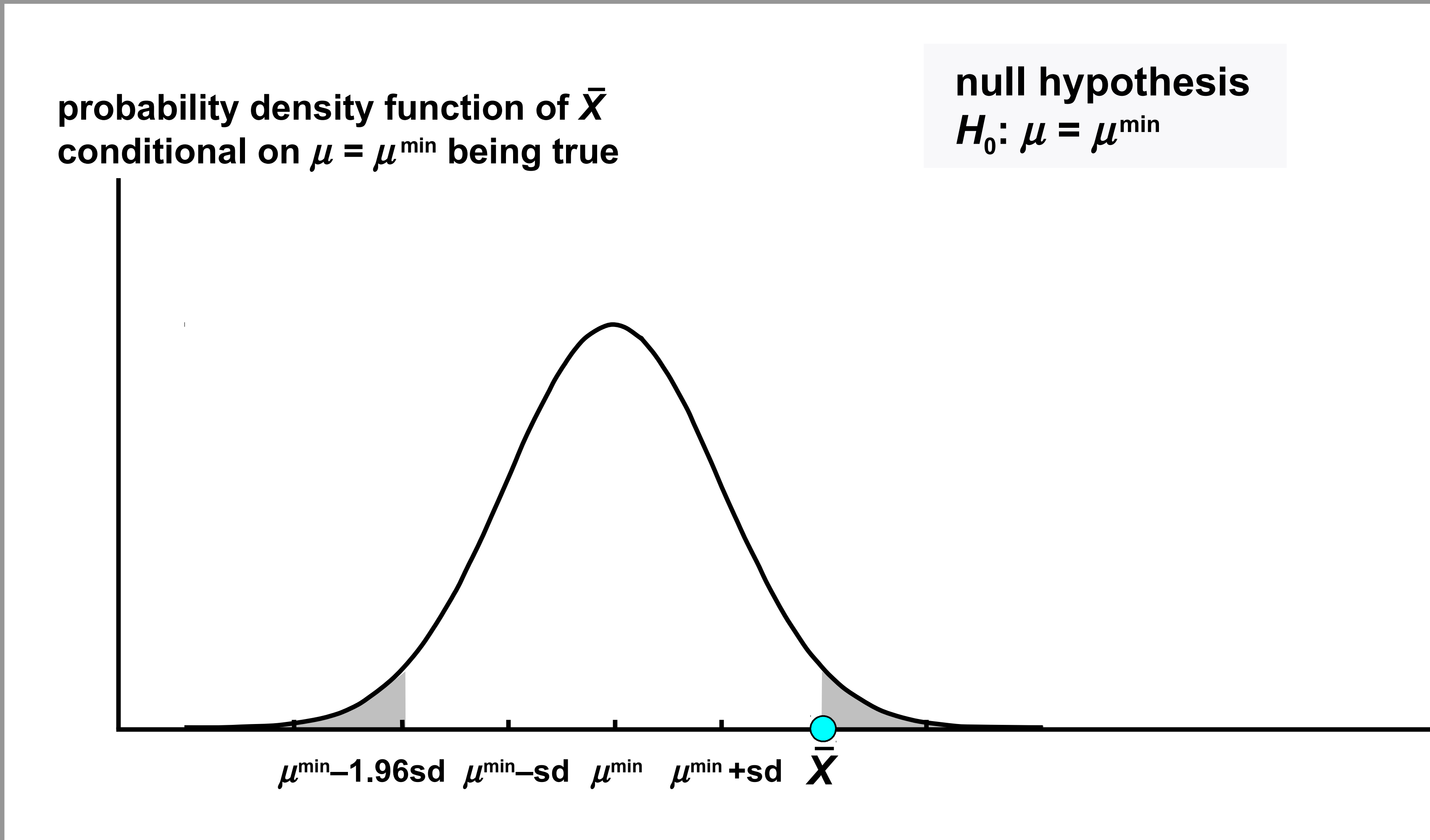
If the hypothetical value of μ were any higher, it would be rejected because $\bar{X}$ would lie inside the left tail of the probability distribution.



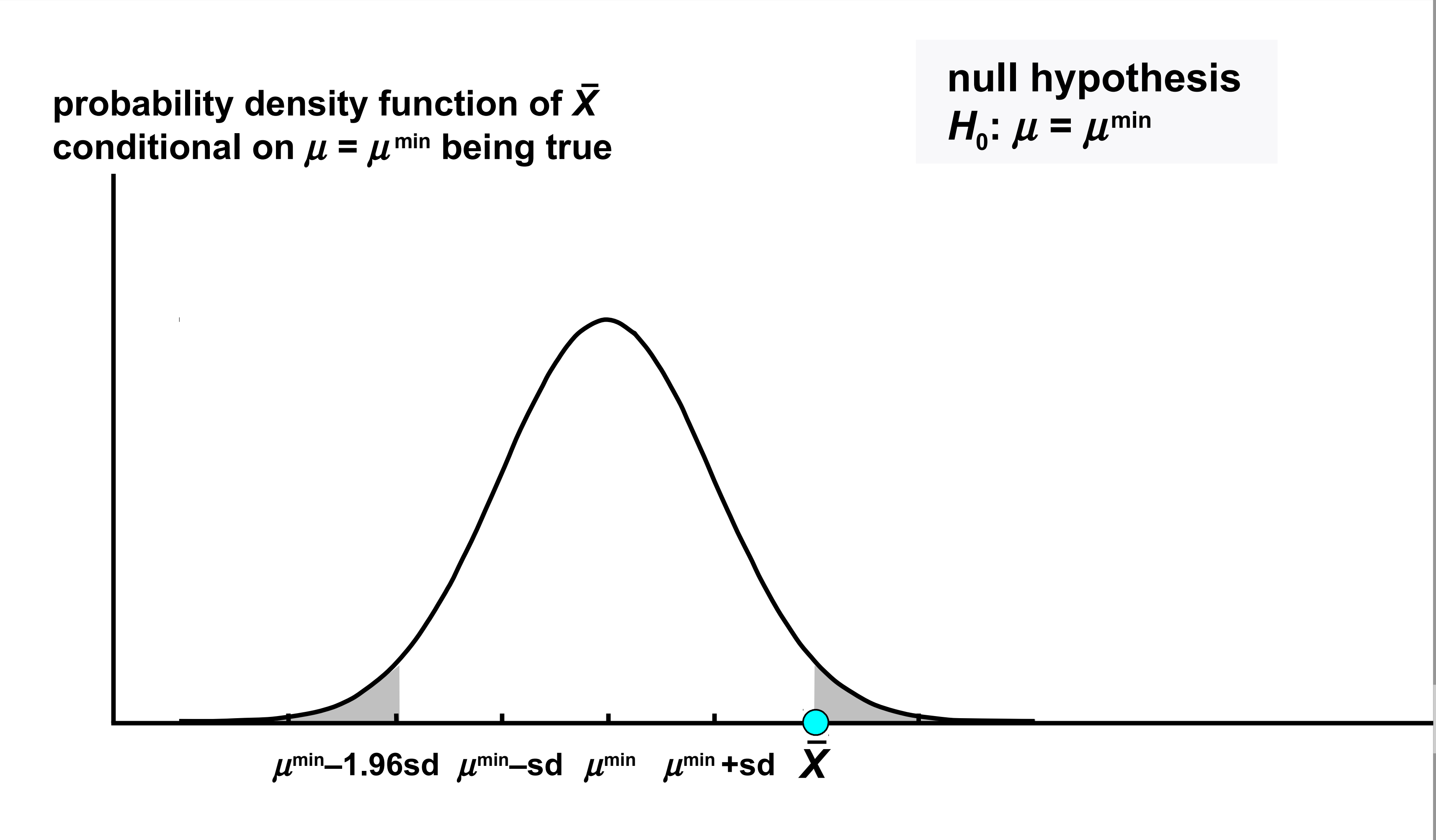
Since $\bar{X}$ lies on the edge of the left 2.5% tail, it must be 1.96 standard deviations less than μ^{max}.



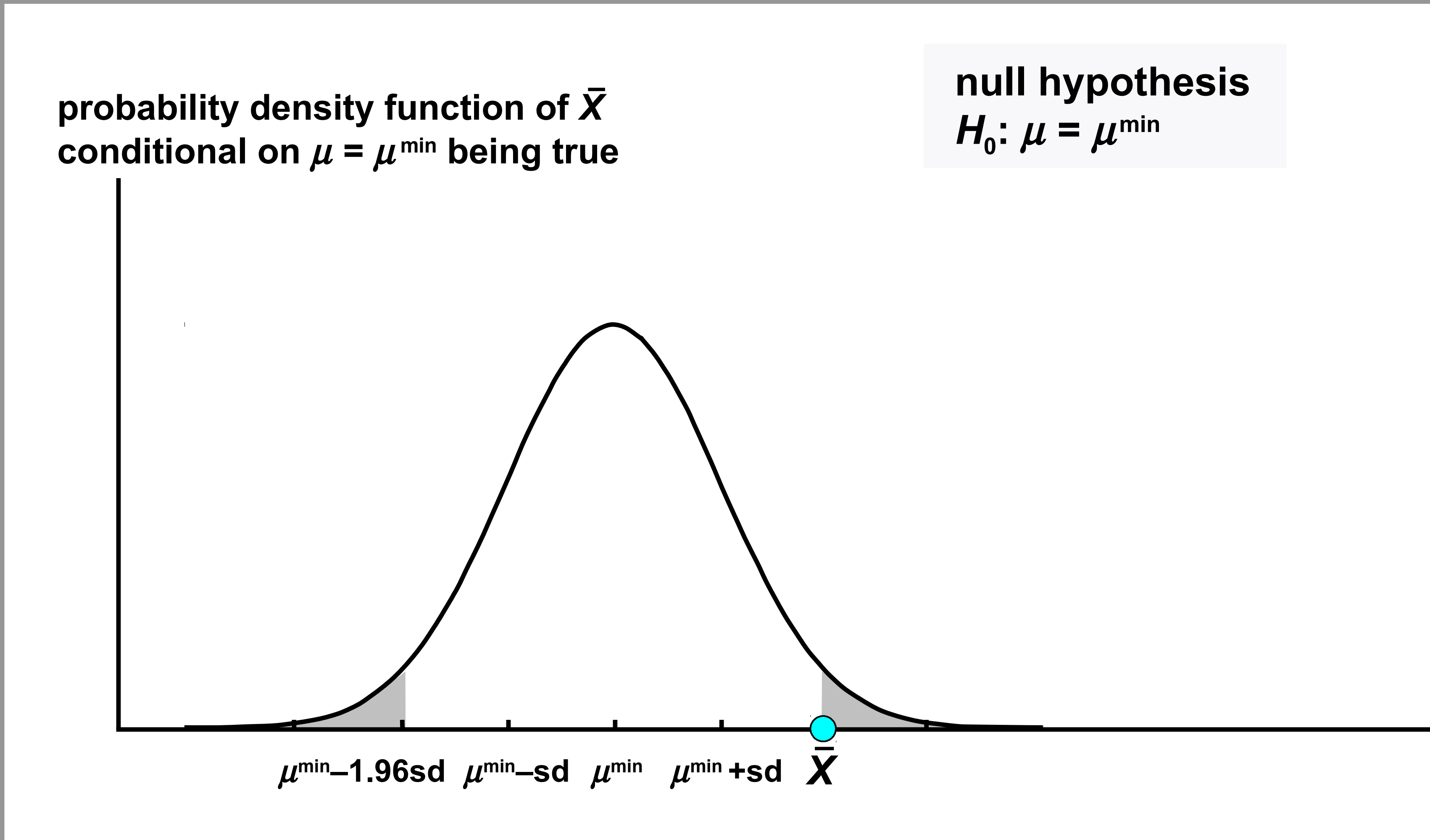
Hence, knowing $\bar{X}$ and the standard deviation, we can calculate μ^{max}.



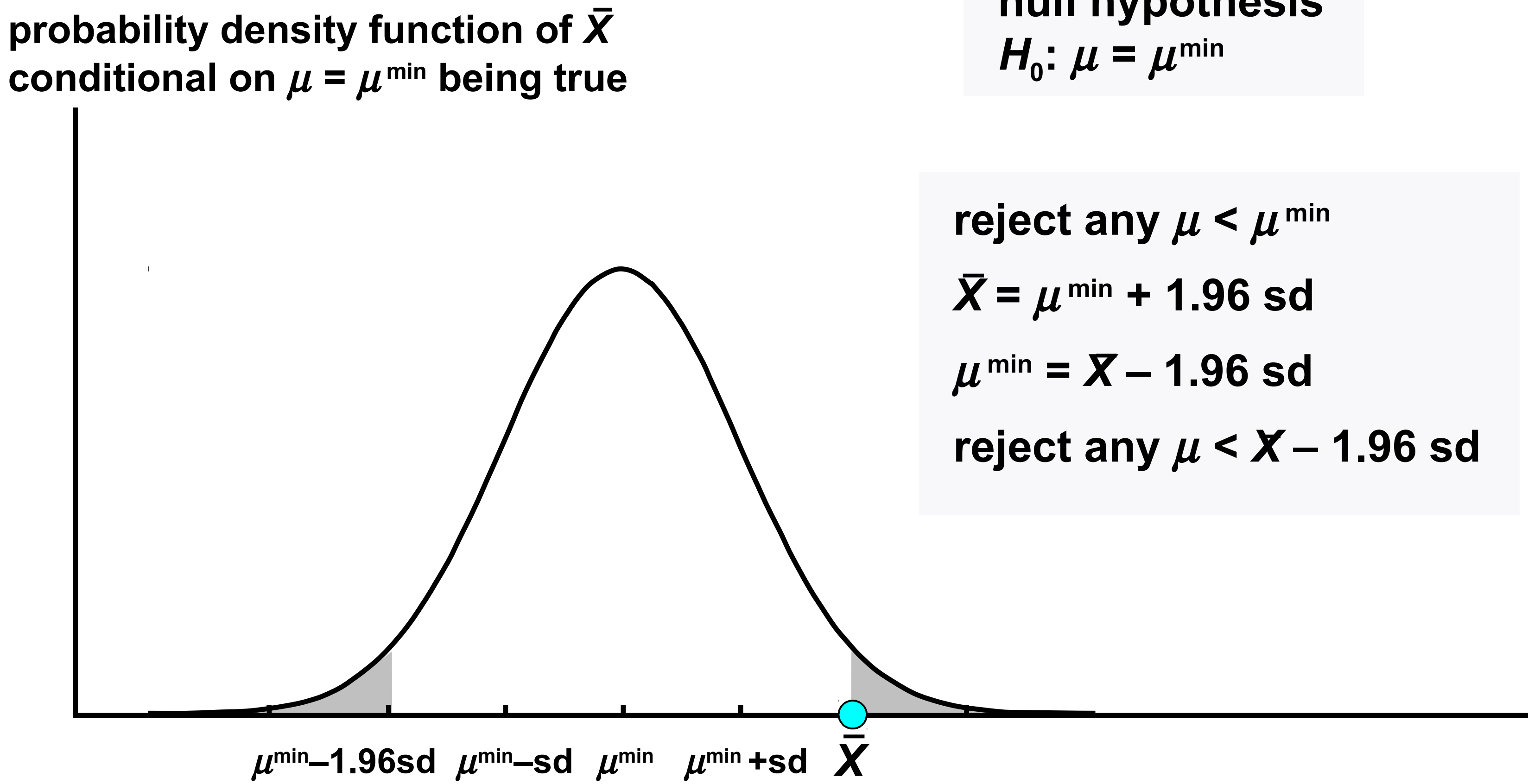
In the same way, we can identify the lowest hypothetical value of μ that is not contradicted by our estimate.



It implies that $\bar{X}$ lies on the edge of the right 2.5% tail of the associated conditional probability distribution. We will call it μ^{min}.



Any lower hypothetical value would be incompatible with our estimate.



null hypothesis
H_0: $\mu = \mu^{min}$

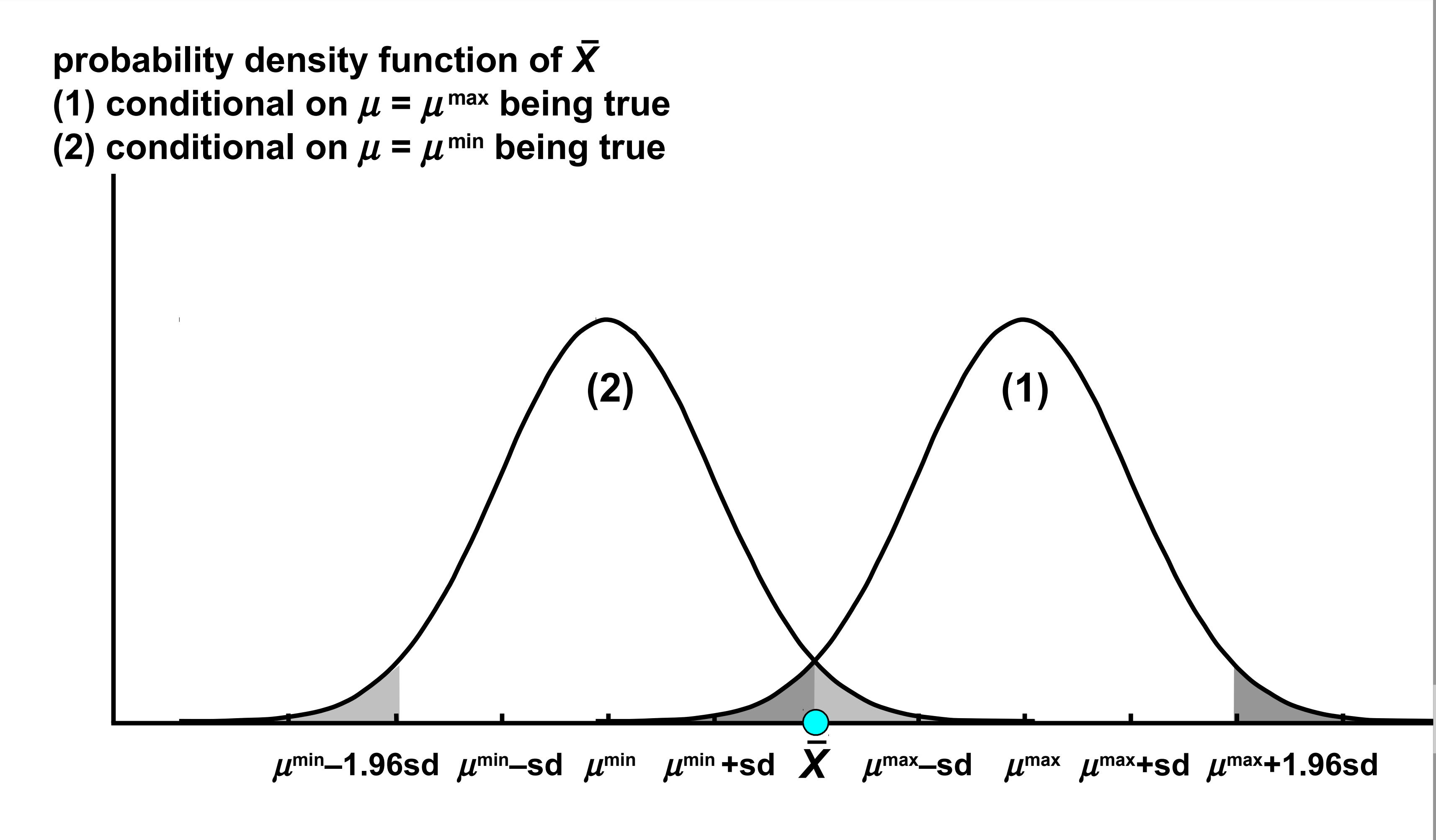
reject any $\mu < \mu^{min}$

$\bar{X} = \mu^{min} + 1.96$ sd

$\mu^{min} = \bar{X} - 1.96$ sd

reject any $\mu < \bar{X} - 1.96$ sd

Since $\bar{X}$ lies on the edge of the right 2.5% tail, $\bar{X}$ is equal to μ^{min} plus 1.96 standard deviations. Hence μ^{min} is equal to $\bar{X}$ minus 1.96 standard deviations.

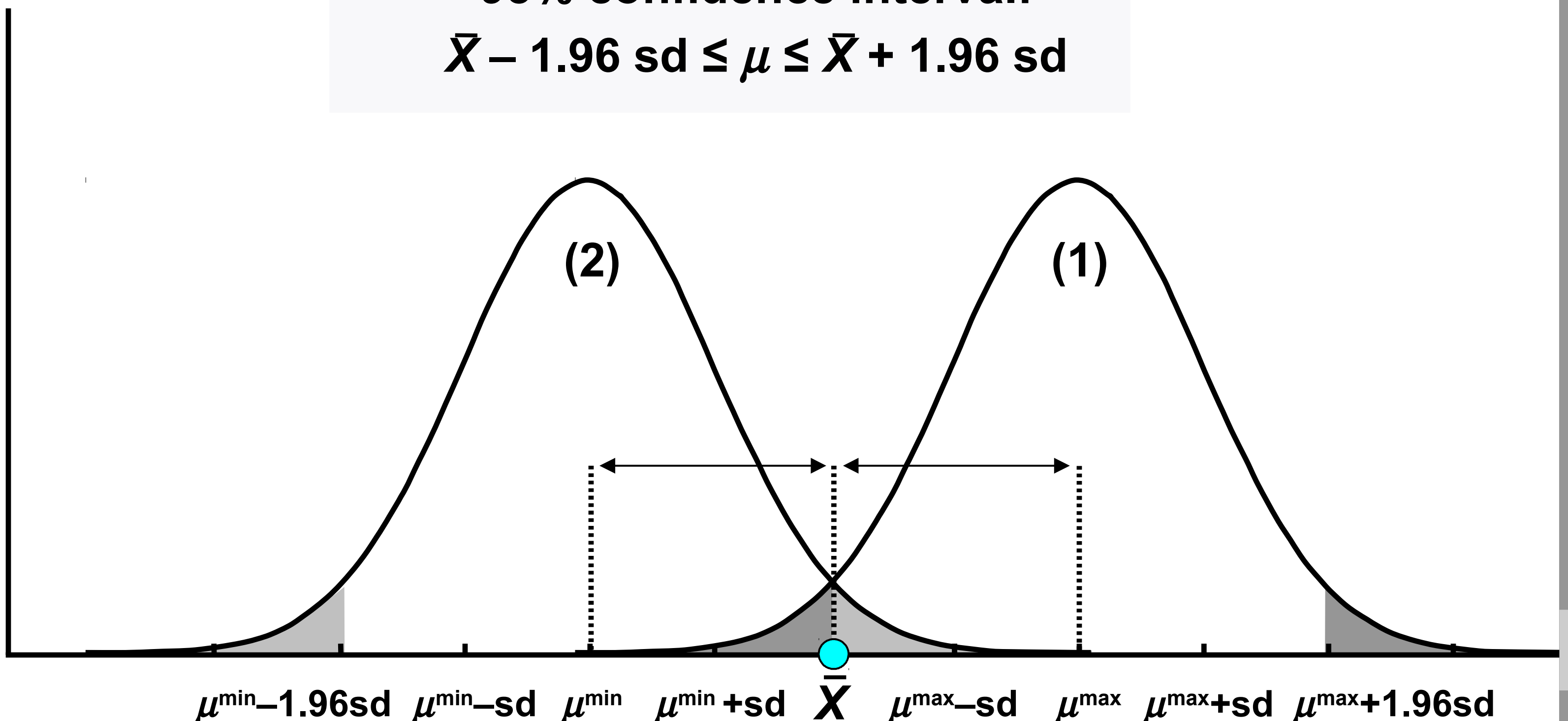


The diagram shows the limiting values of the hypothetical values of μ, together with their associated probability distributions for $\bar{X}$.

CONFIDENCE INTERVALS

reject any $\mu > \mu^{max} = \bar{X} + 1.96$ sd
reject any $\mu < \mu^{min} = \bar{X} - 1.96$ sd

95% confidence interval:
$\bar{X} - 1.96 \text{ sd} \leq \mu \leq \bar{X} + 1.96 \text{ sd}$



Any hypothesis lying in the interval from μ^{min} to μ^{max} would be compatible with the sample estimate (not be rejected by it). We call this interval the 95% confidence interval.

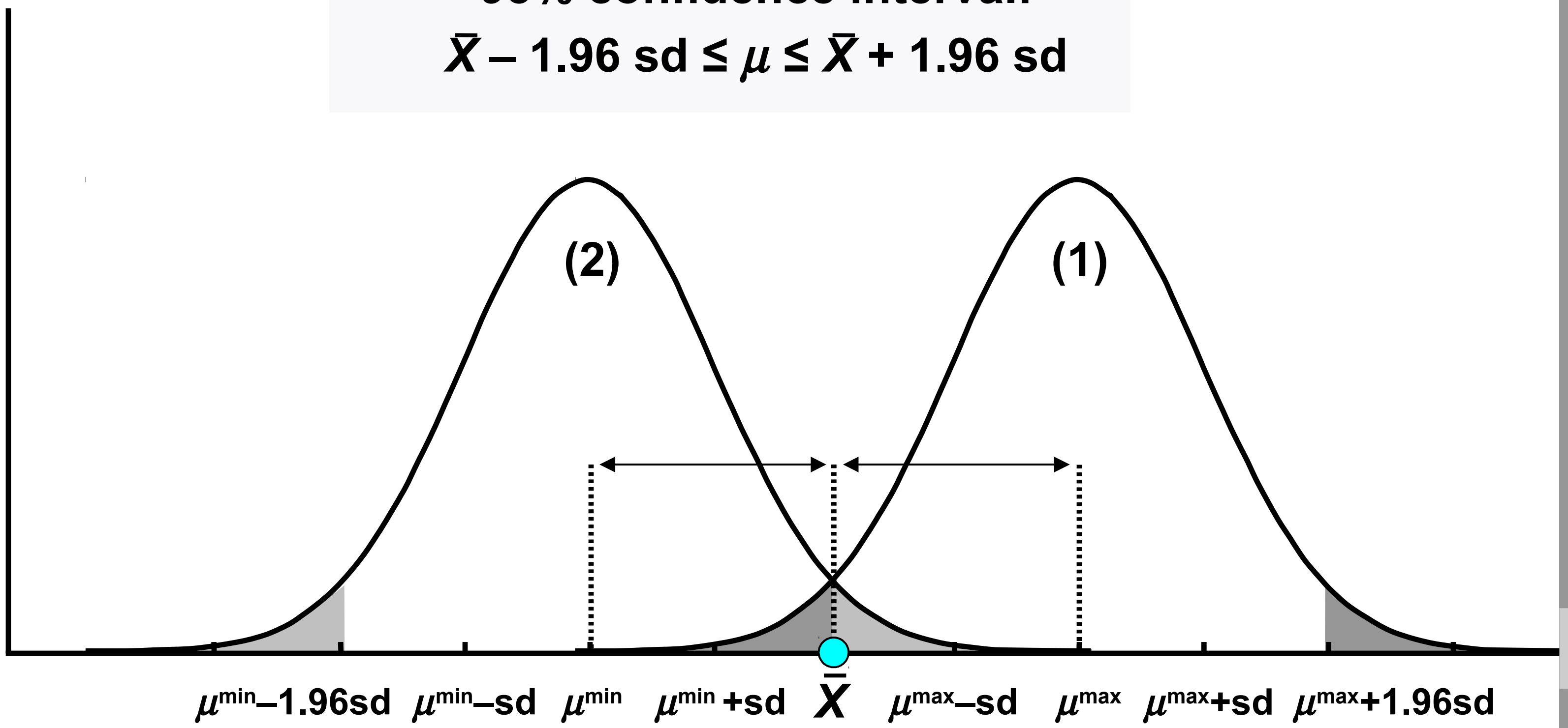
CONFIDENCE INTERVALS

reject any $\mu > \mu^{max} = \bar{X} + 1.96$ sd

reject any $\mu < \mu^{min} = \bar{X} - 1.96$ sd

95% confidence interval:

$\bar{X} - 1.96 \text{ sd} \leq \mu \leq \bar{X} + 1.96 \text{ sd}$



The name arises from a different application of the interval. It can be shown that it will include the true value of the coefficient with 95% probability, provided that the model is correctly specified.

CONFIDENCE INTERVALS

Standard deviation known

95% confidence interval

$$\bar{X} - 1.96 \text{ sd} \le \mu \le \bar{X} + 1.96 \text{ sd}$$

99% confidence interval

$$\bar{X} - 2.58 \text{ sd} \le \mu \le \bar{X} + 2.58 \text{ sd}$$

In exactly the same way, using a 1% significance test to identify hypotheses compatible with our sample estimate, we can construct a 99% confidence interval.

Standard deviation known

95% confidence interval

$$\bar{X} - 1.96 \text{ sd} \leq \mu \leq \bar{X} + 1.96 \text{ sd}$$

99% confidence interval

$$\bar{X} - 2.58 \text{ sd} \leq \mu \leq \bar{X} + 2.58 \text{ sd}$$

μ^{min} and μ^{max} will now be 2.58 standard deviations to the left and to the right of $\bar{X}$, respectively.

CONFIDENCE INTERVALS

Standard deviation known

95% confidence interval

$$\bar{X} - 1.96\ \text{sd} \leq \mu \leq \bar{X} + 1.96\ \text{sd}$$

99% confidence interval

$$\bar{X} - 2.58\ \text{sd} \leq \mu \leq \bar{X} + 2.58\ \text{sd}$$

Standard deviation estimated by standard error

95% confidence interval

$$\bar{X} - t_{\text{crit}\ (5\%)}\ \text{se} \leq \mu \leq \bar{X} + t_{\text{crit}\ (5\%)}\ \text{se}$$

99% confidence interval

$$\bar{X} - t_{\text{crit}\ (1\%)}\ \text{se} \leq \mu \leq \bar{X} + t_{\text{crit}\ (1\%)}\ \text{se}$$

Until now we have assumed that we know the standard deviation of the distribution. In practice we have to estimate it.

CONFIDENCE INTERVALS

Standard deviation known

95% confidence interval

$$\bar{X} - 1.96 \text{ sd} \le \mu \le \bar{X} + 1.96 \text{ sd}$$

99% confidence interval

$$\bar{X} - 2.58 \text{ sd} \le \mu \le \bar{X} + 2.58 \text{ sd}$$

Standard deviation estimated by standard error

95% confidence interval

$$\bar{X} - t_{\text{crit (5\%)}} \text{ se} \le \mu \le \bar{X} + t_{\text{crit (5\%)}} \text{ se}$$

99% confidence interval

$$\bar{X} - t_{\text{crit (1\%)}} \text{ se} \le \mu \le \bar{X} + t_{\text{crit (1\%)}} \text{ se}$$

As a consequence, the *t* distribution has to be used instead of the normal distribution when locating μ^{min} and μ^{max}.

CONFIDENCE INTERVALS

Standard deviation known

95% confidence interval

$$\bar{X} - 1.96 \text{ sd} \leq \mu \leq \bar{X} + 1.96 \text{ sd}$$

99% confidence interval

$$\bar{X} - 2.58 \text{ sd} \leq \mu \leq \bar{X} + 2.58 \text{ sd}$$

Standard deviation estimated by standard error

95% confidence interval

$$\bar{X} - t_{\text{crit (5\%)}} \text{ se} \leq \mu \leq \bar{X} + t_{\text{crit (5\%)}} \text{ se}$$

99% confidence interval

$$\bar{X} - t_{\text{crit (1\%)}} \text{ se} \leq \mu \leq \bar{X} + t_{\text{crit (1\%)}} \text{ se}$$

This implies that the standard error should be multiplied by the critical value of t, given the significance level and number of degrees of freedom, when determining the limits of the interval.

The content of this slideshow comes from Section R.12 of C. Dougherty, *Introduction to Econometrics*, fifth edition 2016, Oxford University Press. Additional (free) resources for both students and instructors may be downloaded from the OUP Online Resource Centre
http://www.oxfordtextbooks.co.uk/orc/dougherty5e/

Individuals studying econometrics on their own who feel that they might benefit from participation in a formal course should consider the London School of Economics summer school course
EC212 Introduction to Econometrics
http://www2.lse.ac.uk/study/summerSchools/summerSchool/Home.aspx
or the University of London International Programmes distance learning course
EC2020 Elements of Econometrics
www.londoninternational.ac.uk/lse.