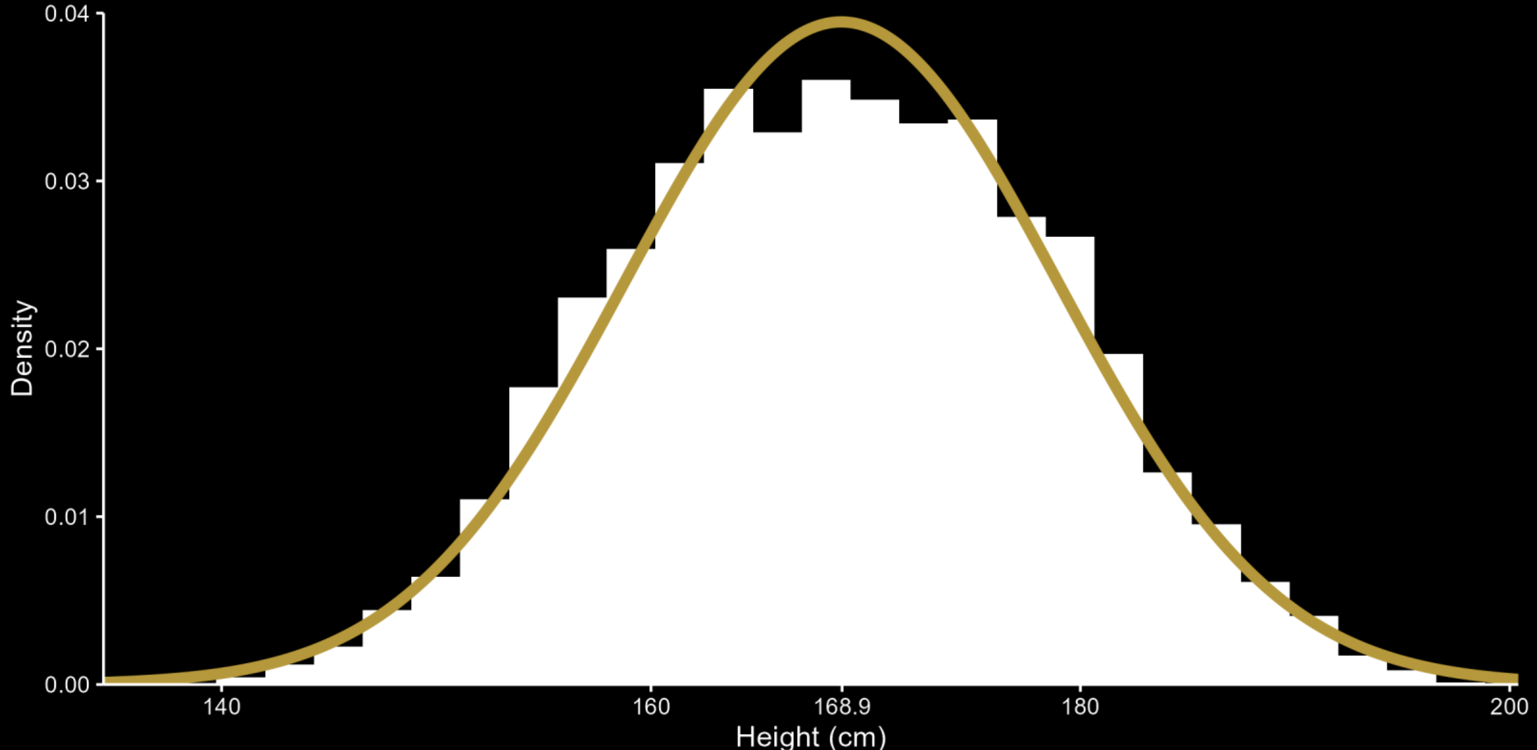
Central limit theorem

NHH TECH3

Why is height normally distributed?



THE CENTRAL LIMIT THEOREM

Let $Y_1, \ldots, Y_n$ be a sequence of independent and identically distributed random variables having a distribution with expectation μ and finite variance σ^2. Then

$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow[n \to \infty]{\text{Distribution}} N\left(\mu, \ \frac{\sigma^2}{n}\right)$$

NHH
TECH3

Consequences

- Sampling distribution of the sample mean is asymptotically Gaussian
- Note that we do not assume any specific distribution of the Ys!
- The standard deviation of this Gaussian distribution can be estimated as $s_n/\sqrt{n}$
- As n increases, the standard deviation decreases, and the sample mean approaches the population mean μ.

NHH TECH3

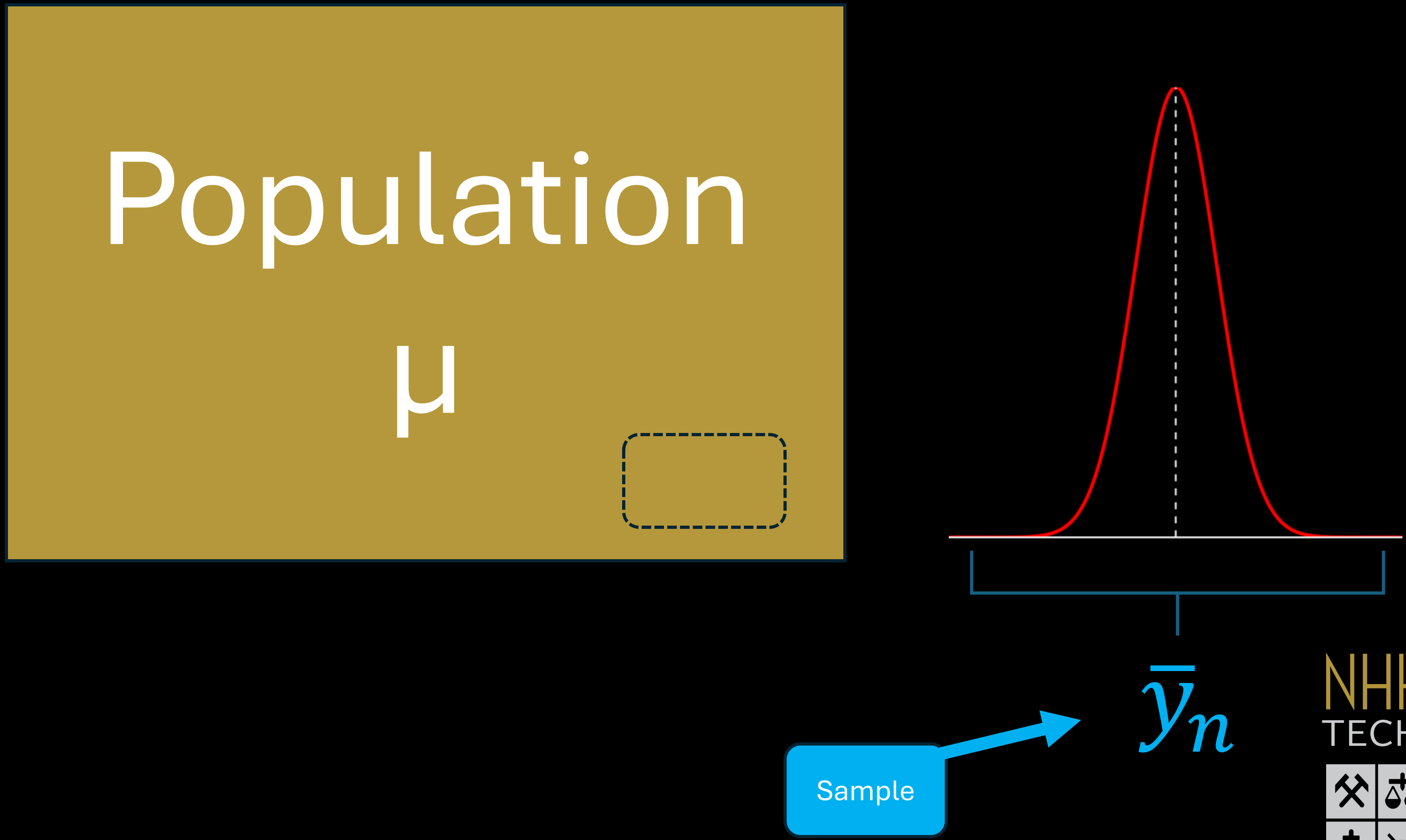
Central limit theorem

- Partly why the normal distribution is so important!
- Useful whenever we deal with a mean (or a sum!)
- Approximation of distributions
- Inference on population parameters
- Hypothesis testing

Population

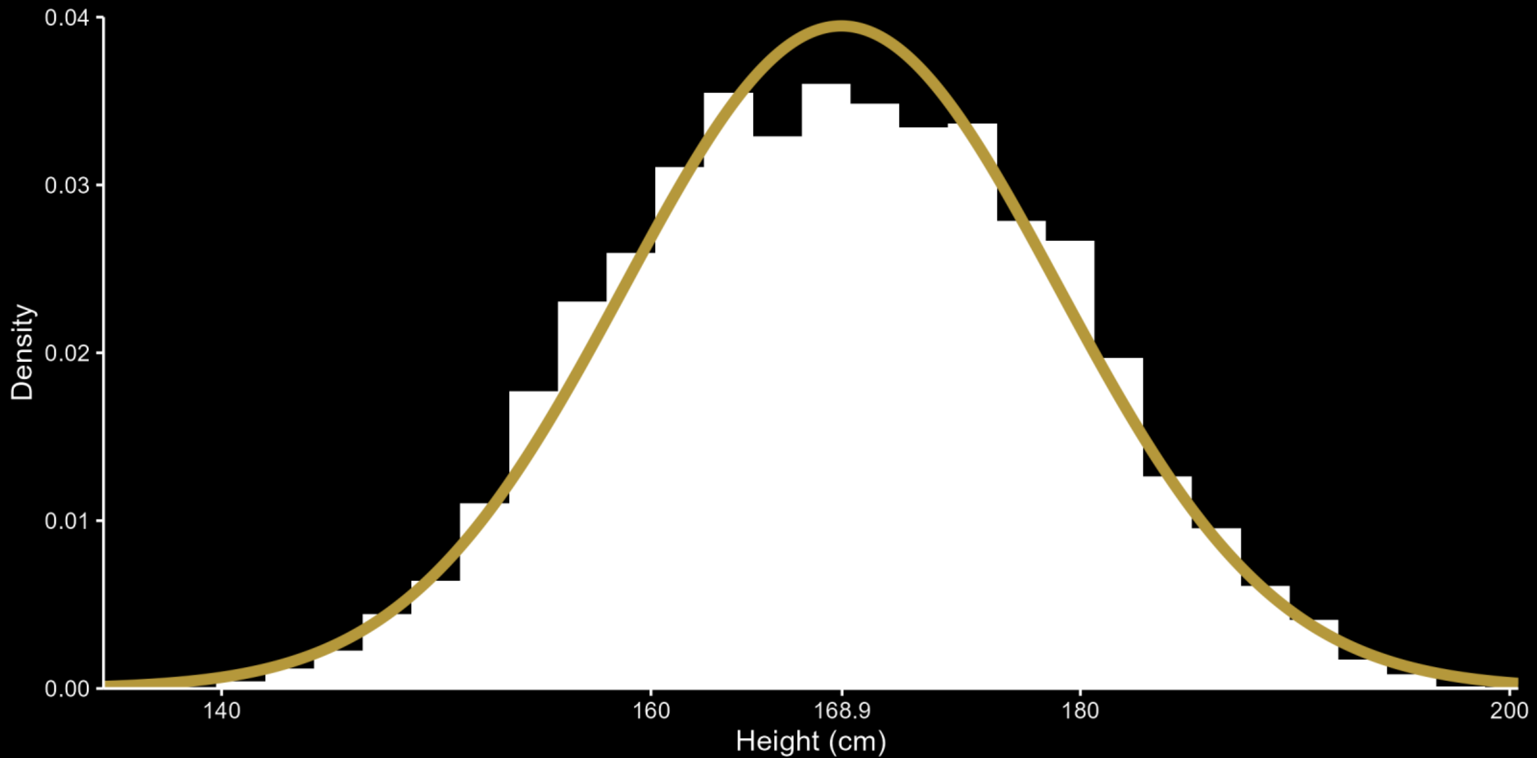
μ

Sample

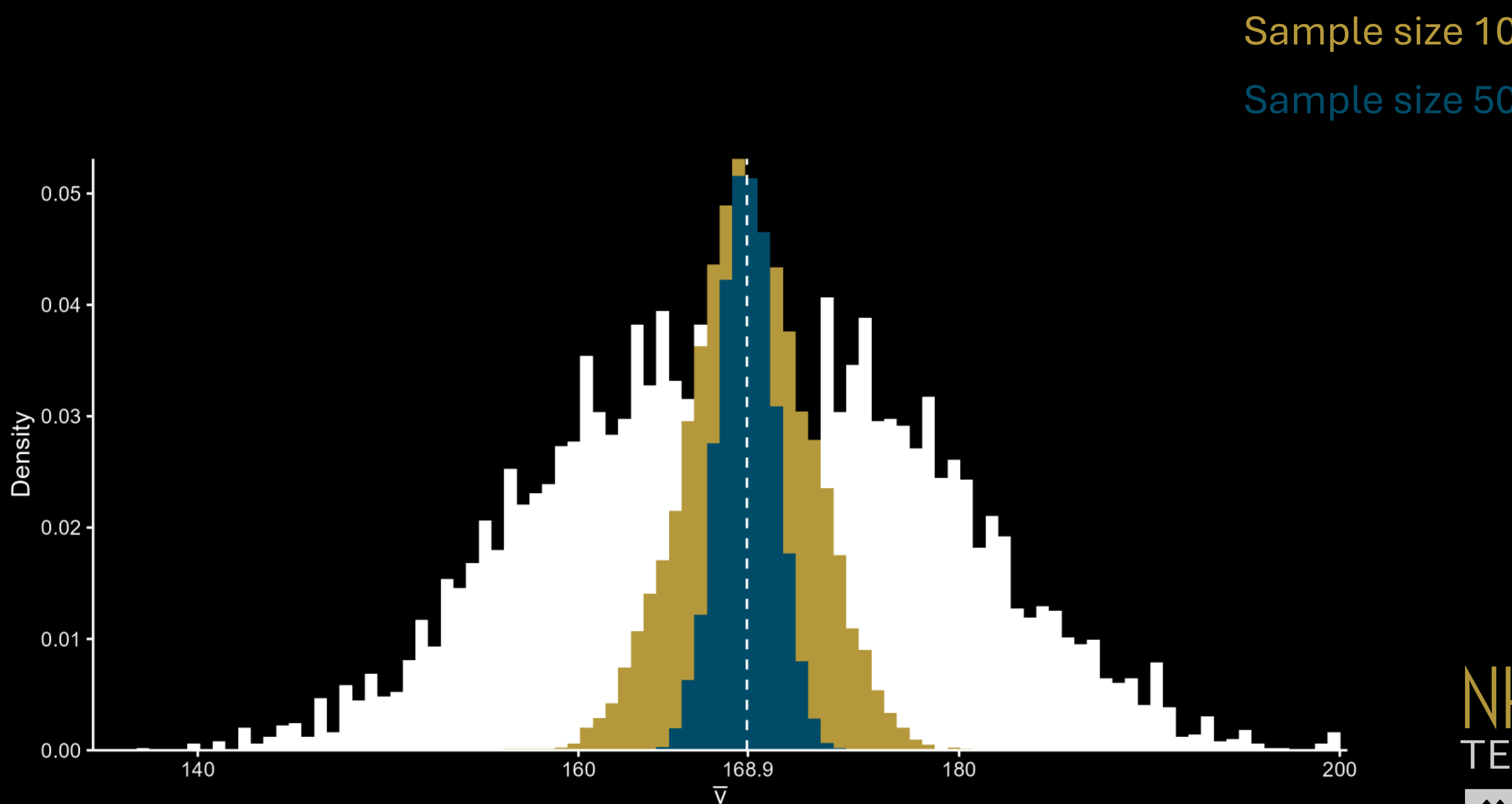


NHH
TECH3

Why is height normally distributed?



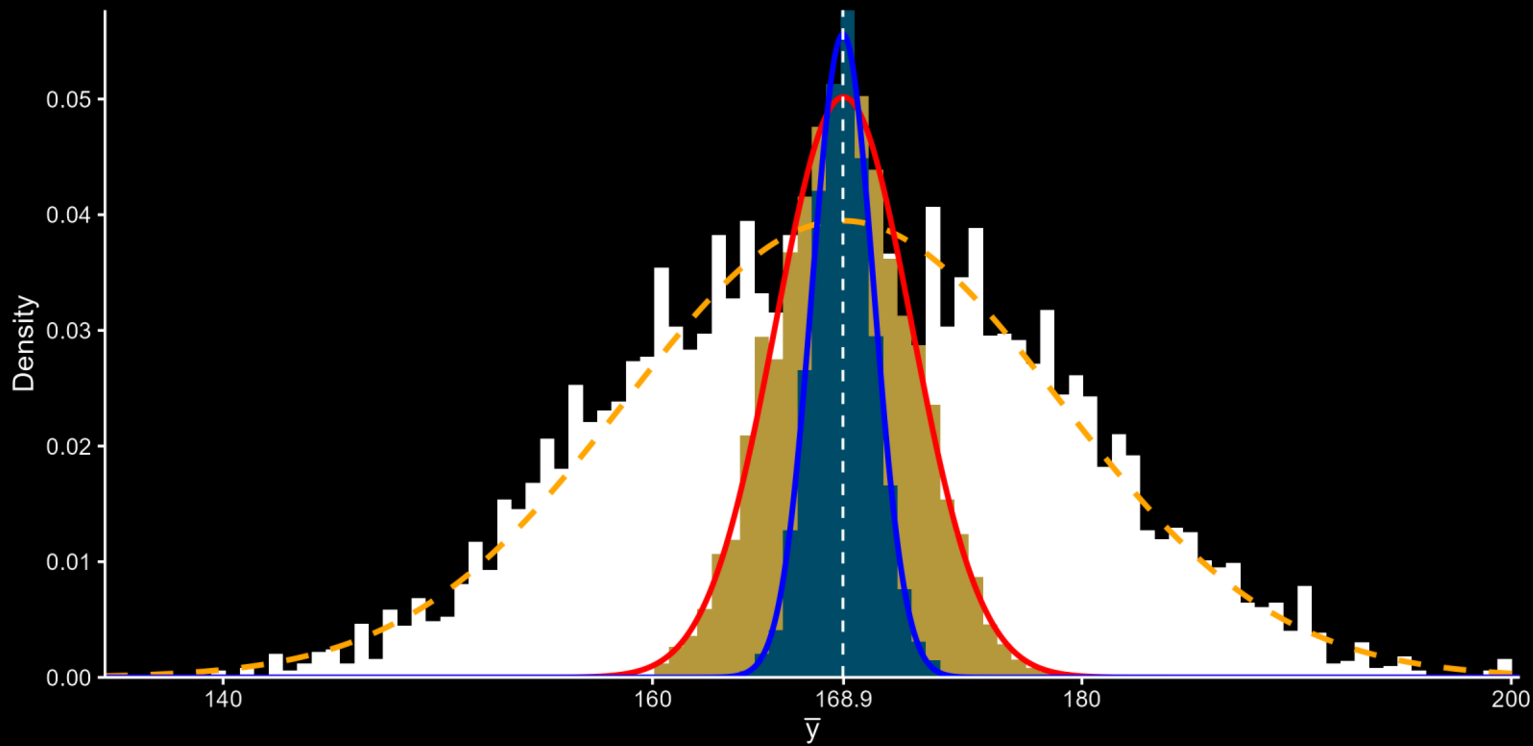
NHH
TECH3

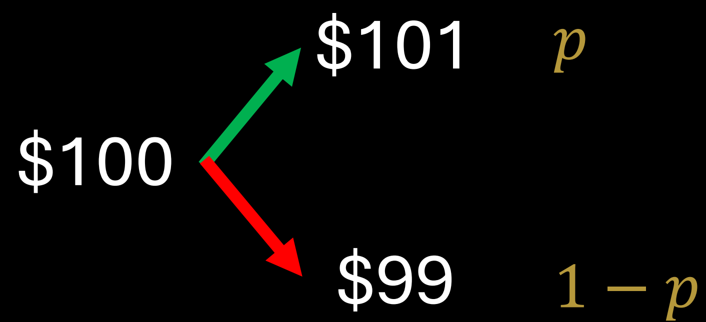

Sample size 10
Sample size 50
Density
0.05
0.04
0.03
0.02
0.01
0.00
140
160
168.9
180
200
ȳ

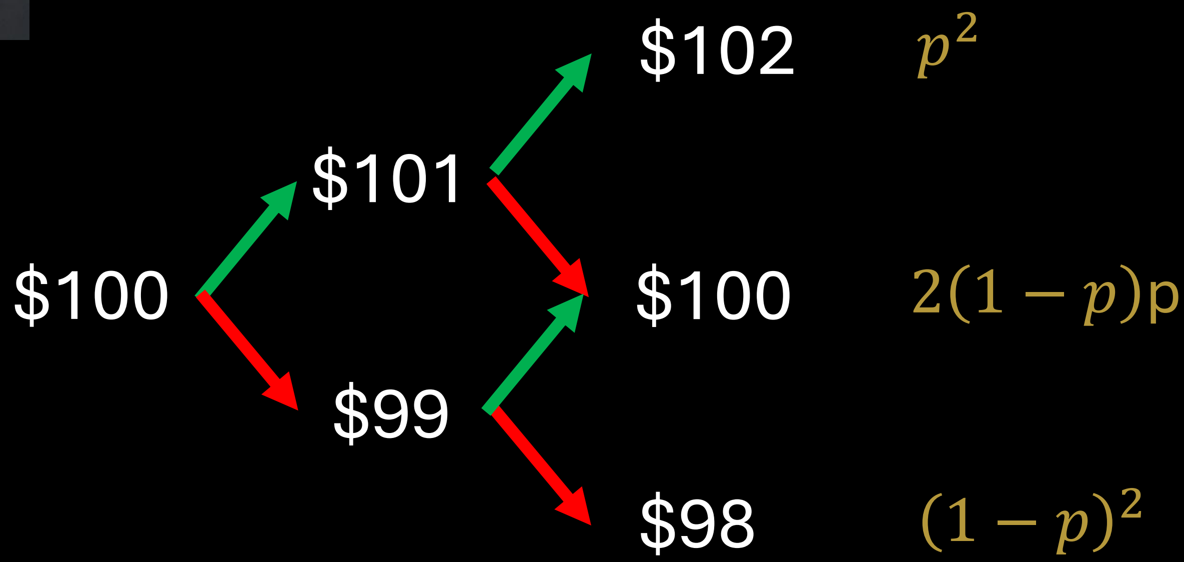
NHH TECH3

Sample size 10

Sample size 50




$100
$101
p
$99
1 − p


$100
$101
$99
$102
$100
$98
p^2
$2(1-p)p$
$(1-p)^2$


$100
$101
$99
$102
$100
$98
$103
$101
$99
$97
$104
$102
$100
$98
$96
$105
$103
$101
$99
$97
$95
n = 1
n = 2
n = 3

$$P(X_i = 1) = p$$

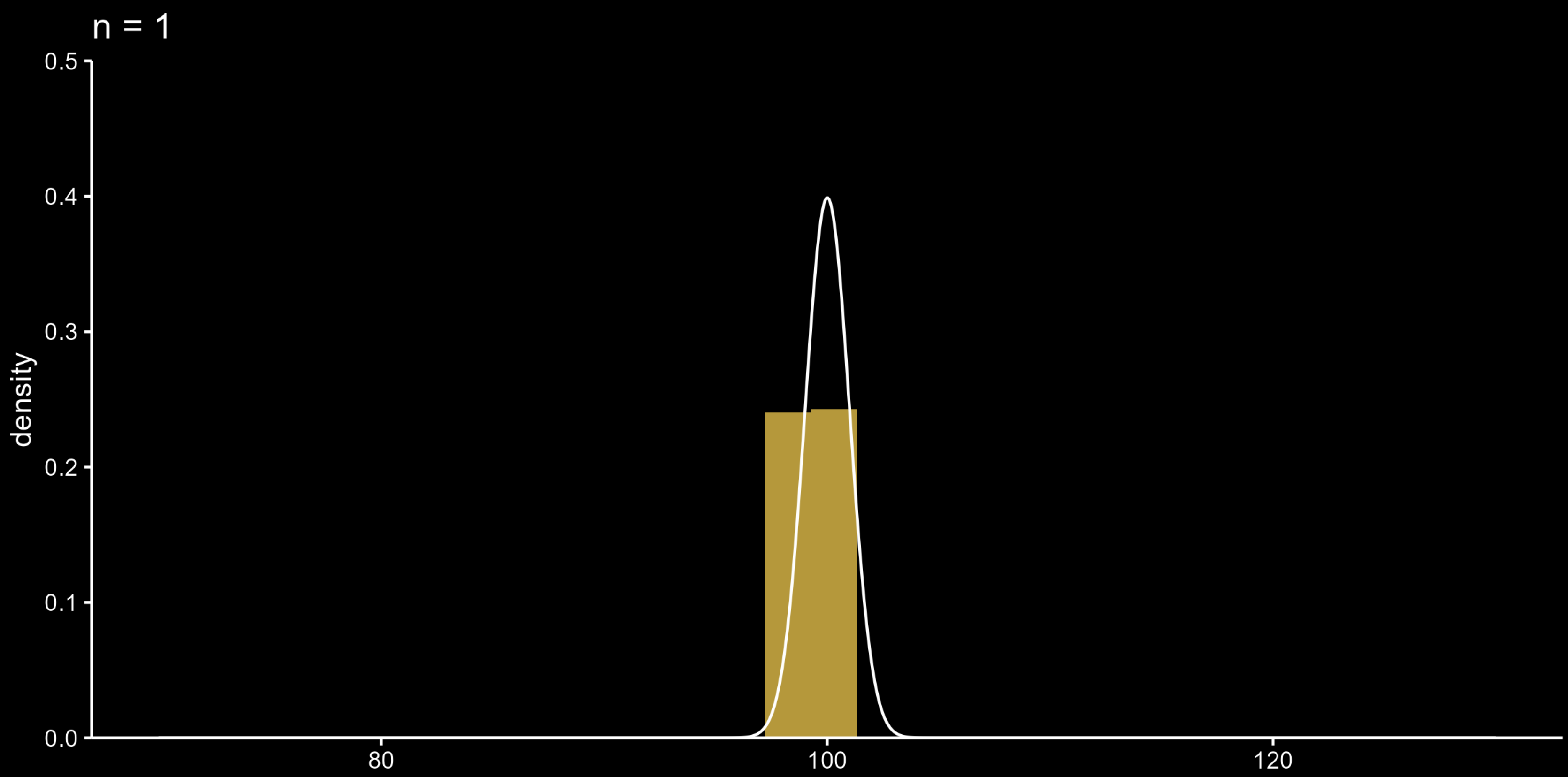
$$P(X_i = -1) = 1 - p$$

$$Y_n = 100 + \sum_{i=1}^{n} X_i$$

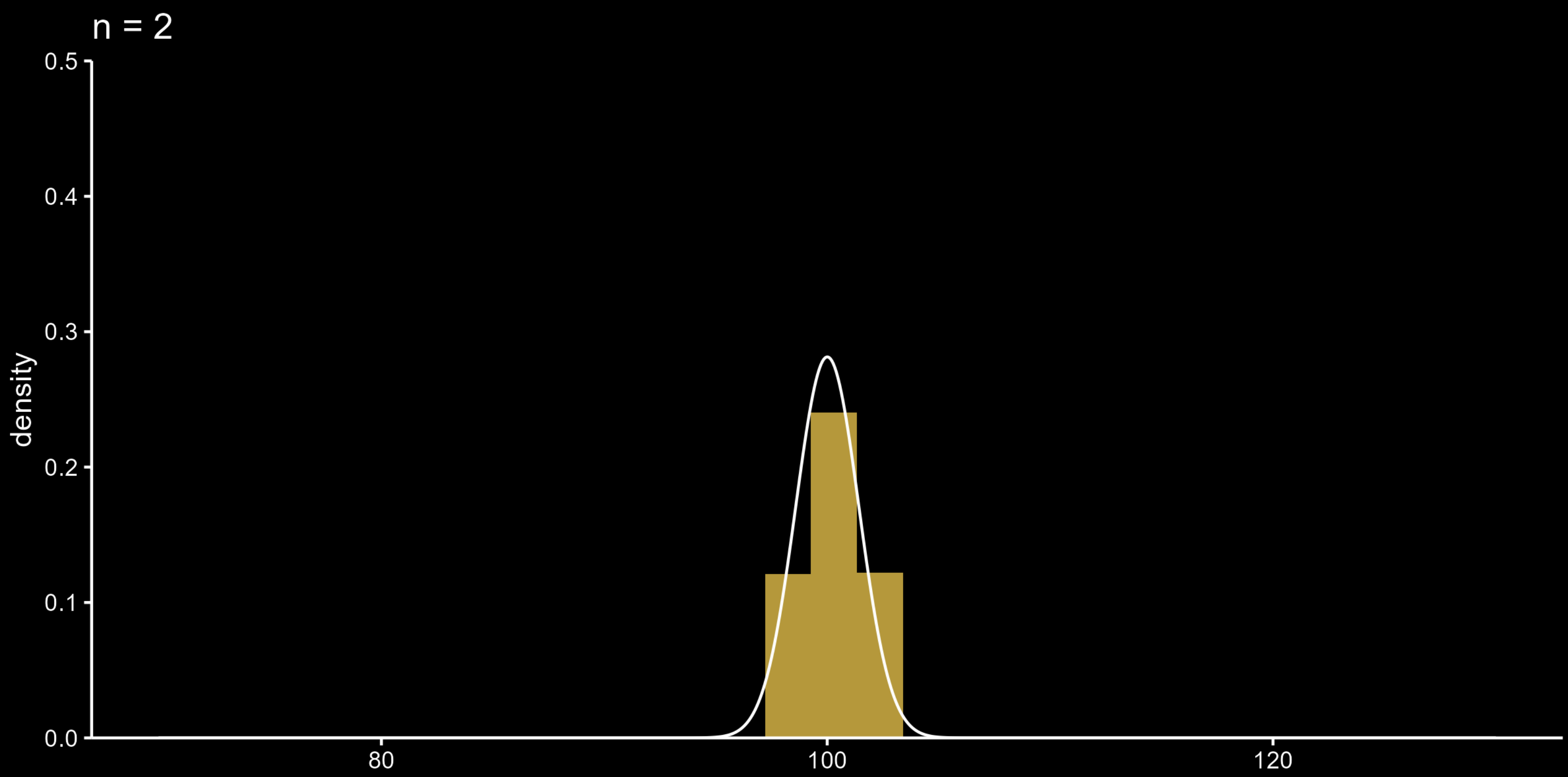
$$E(Y_n) = 100 \cdot (2p - 1)$$

$$\text{Var}(Y_n) = 2p \cdot n$$

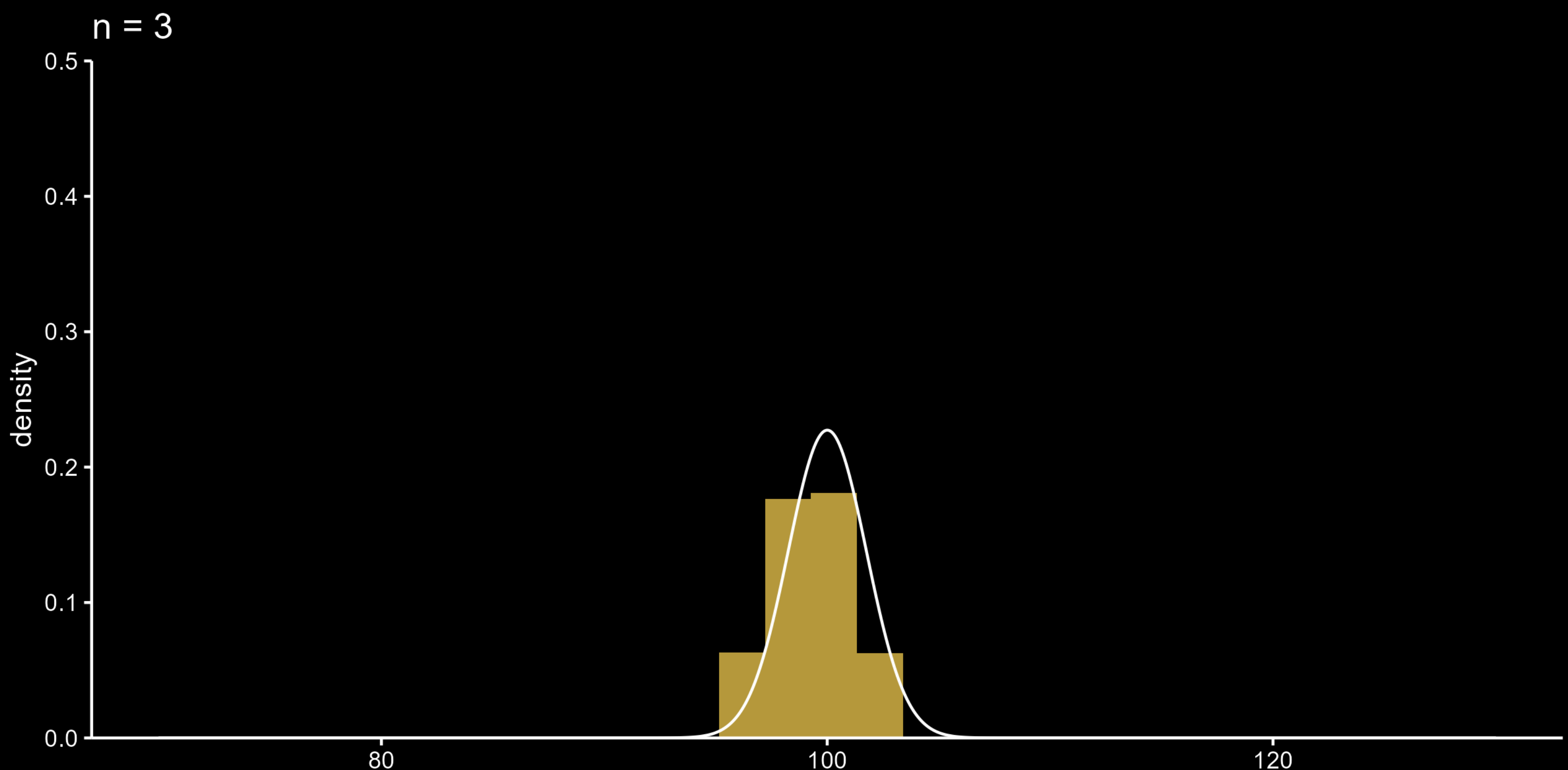
NHH TECH3



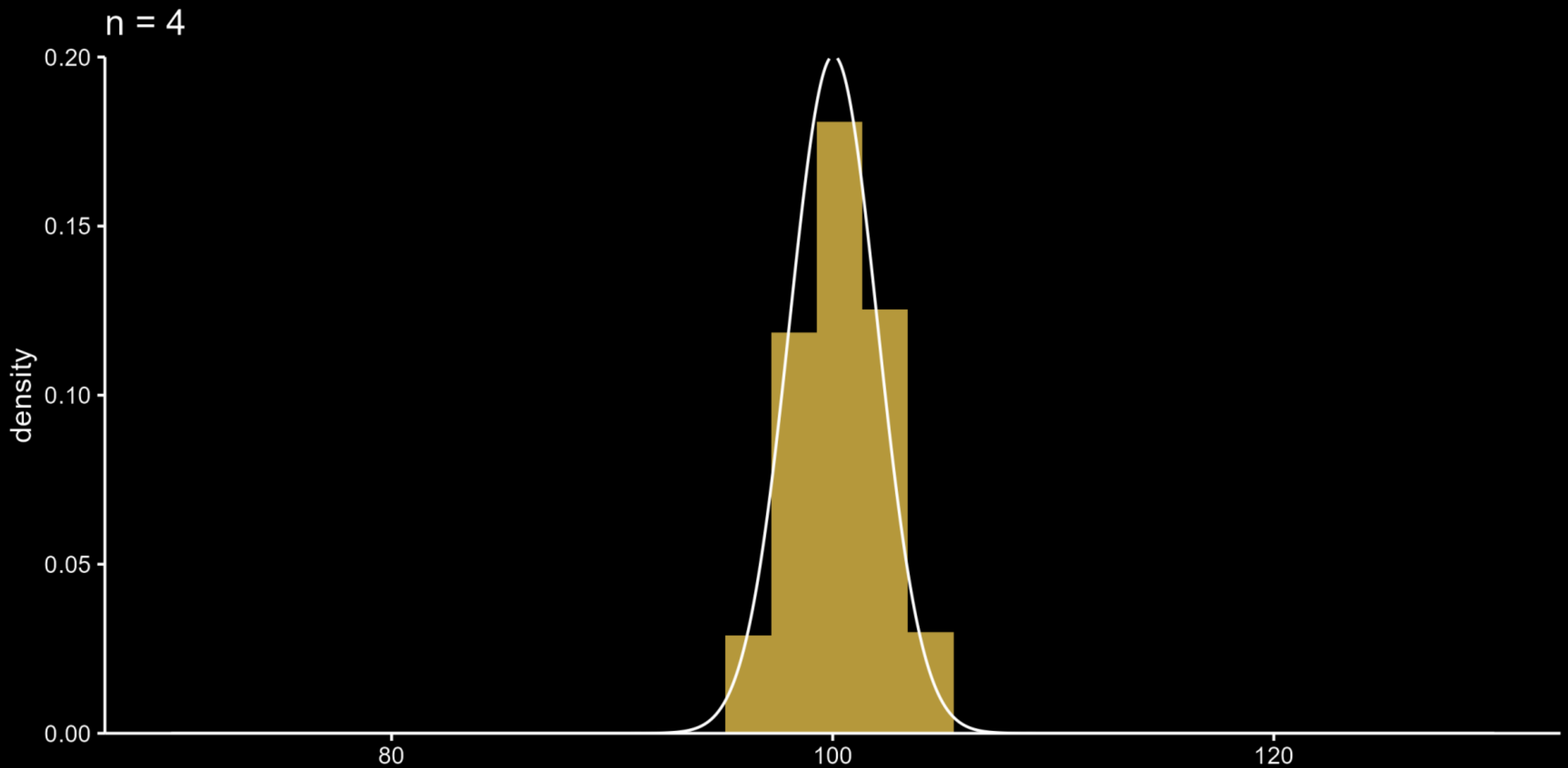
NHH TECH3



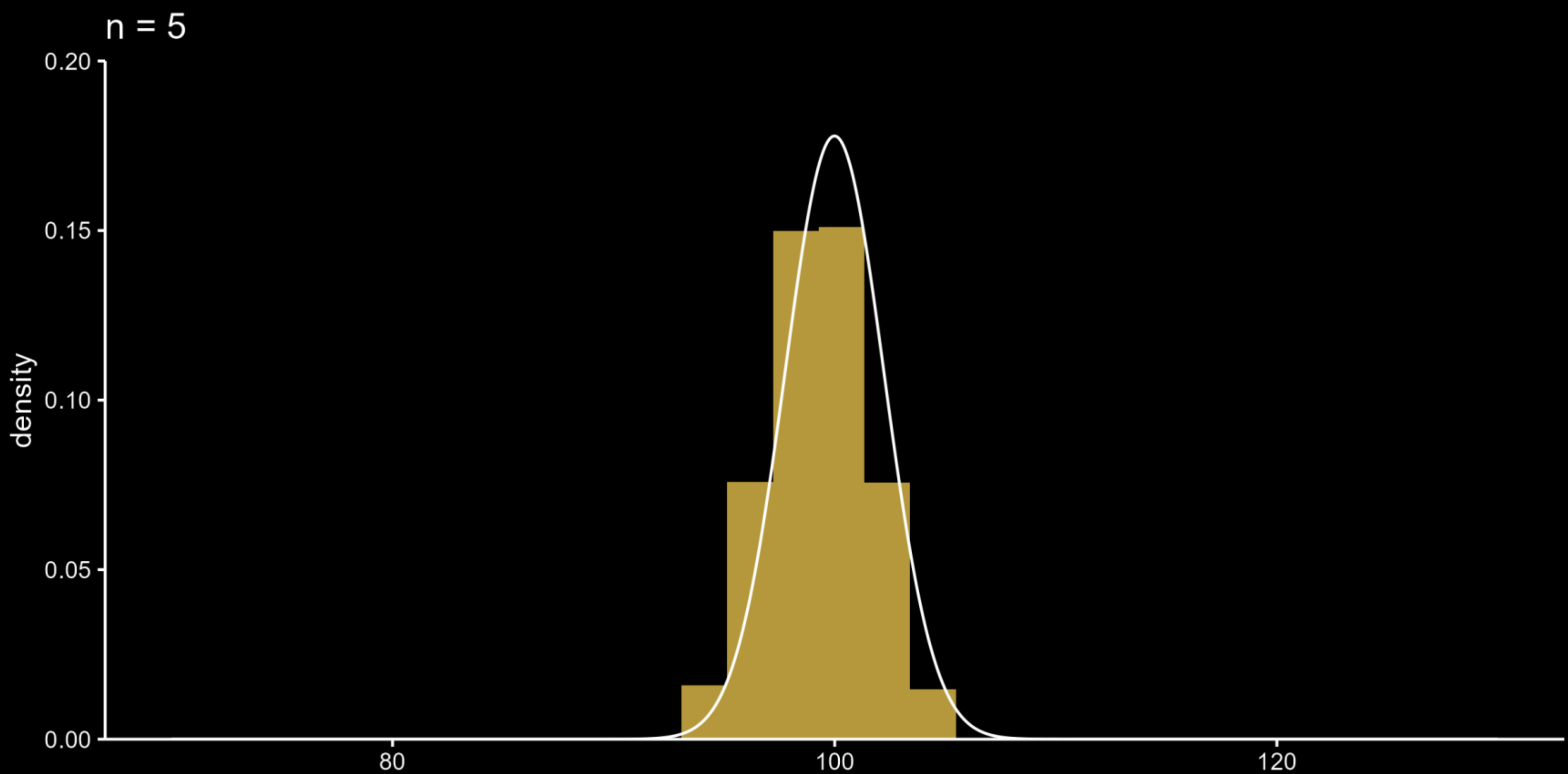
NHH
TECH3

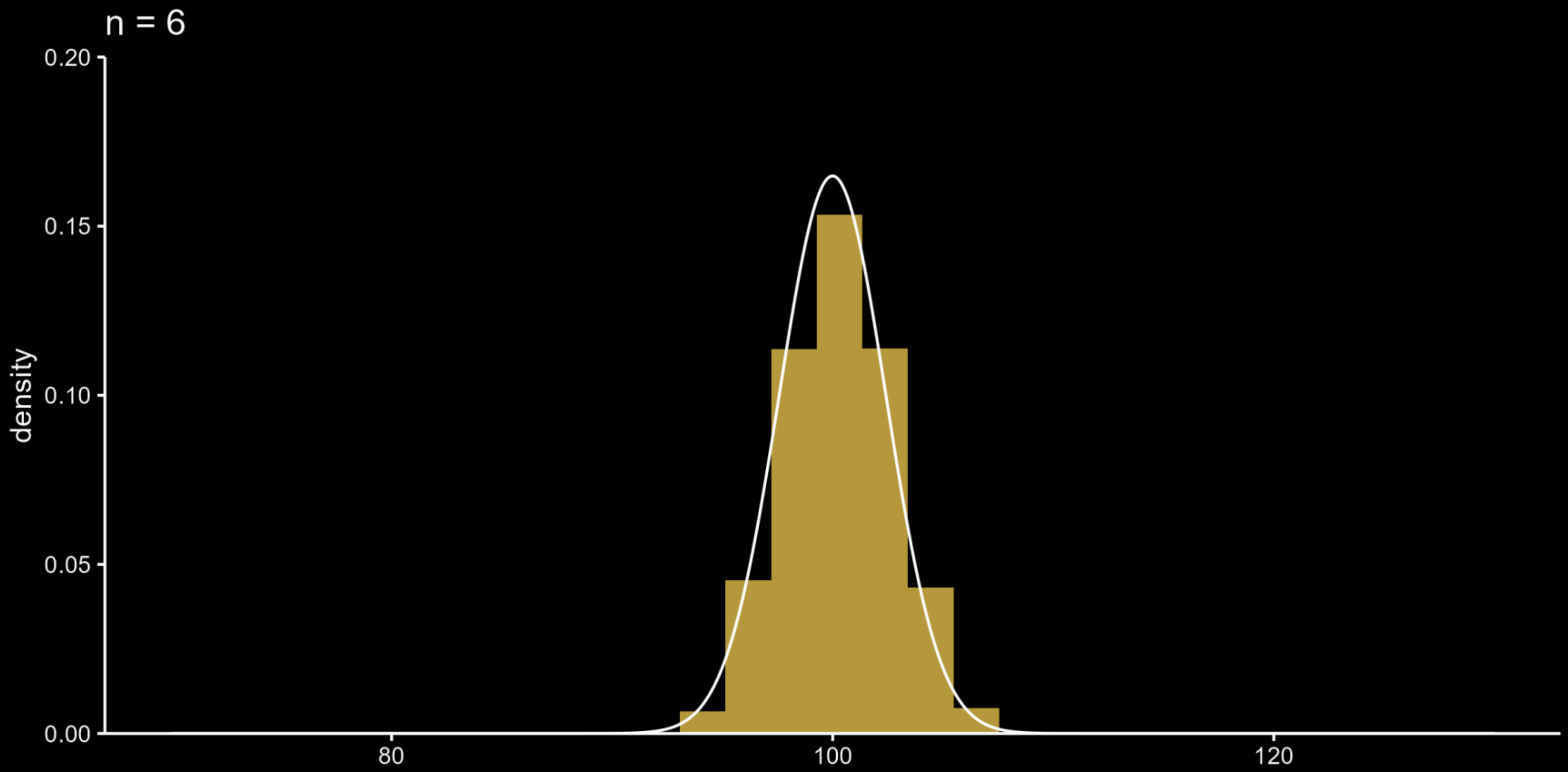


n = 3
density
0.5
0.4
0.3
0.2
0.1
0.0
80
100
120

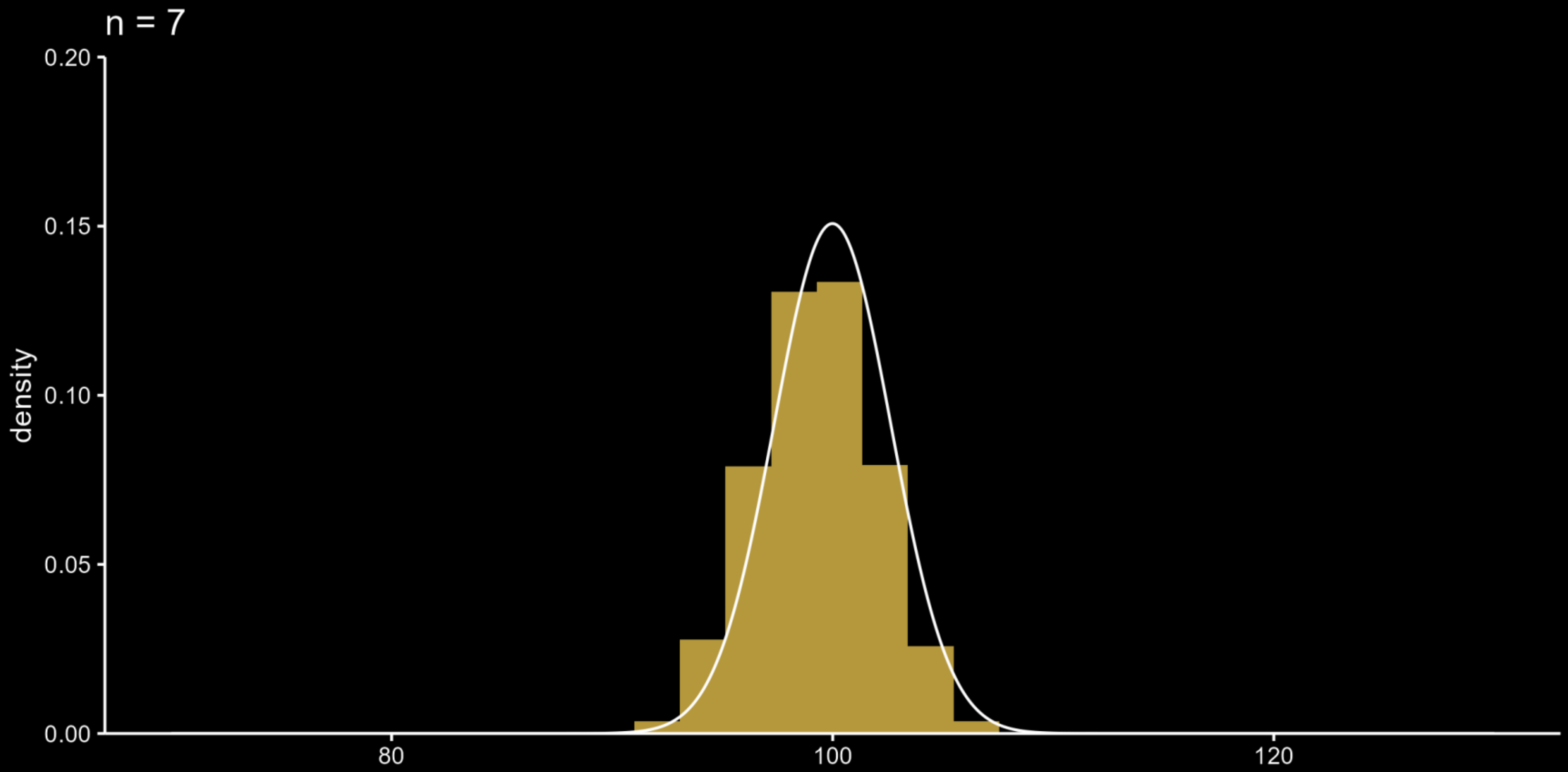


n = 4
density
0.20
0.15
0.10
0.05
0.00
80
100
120

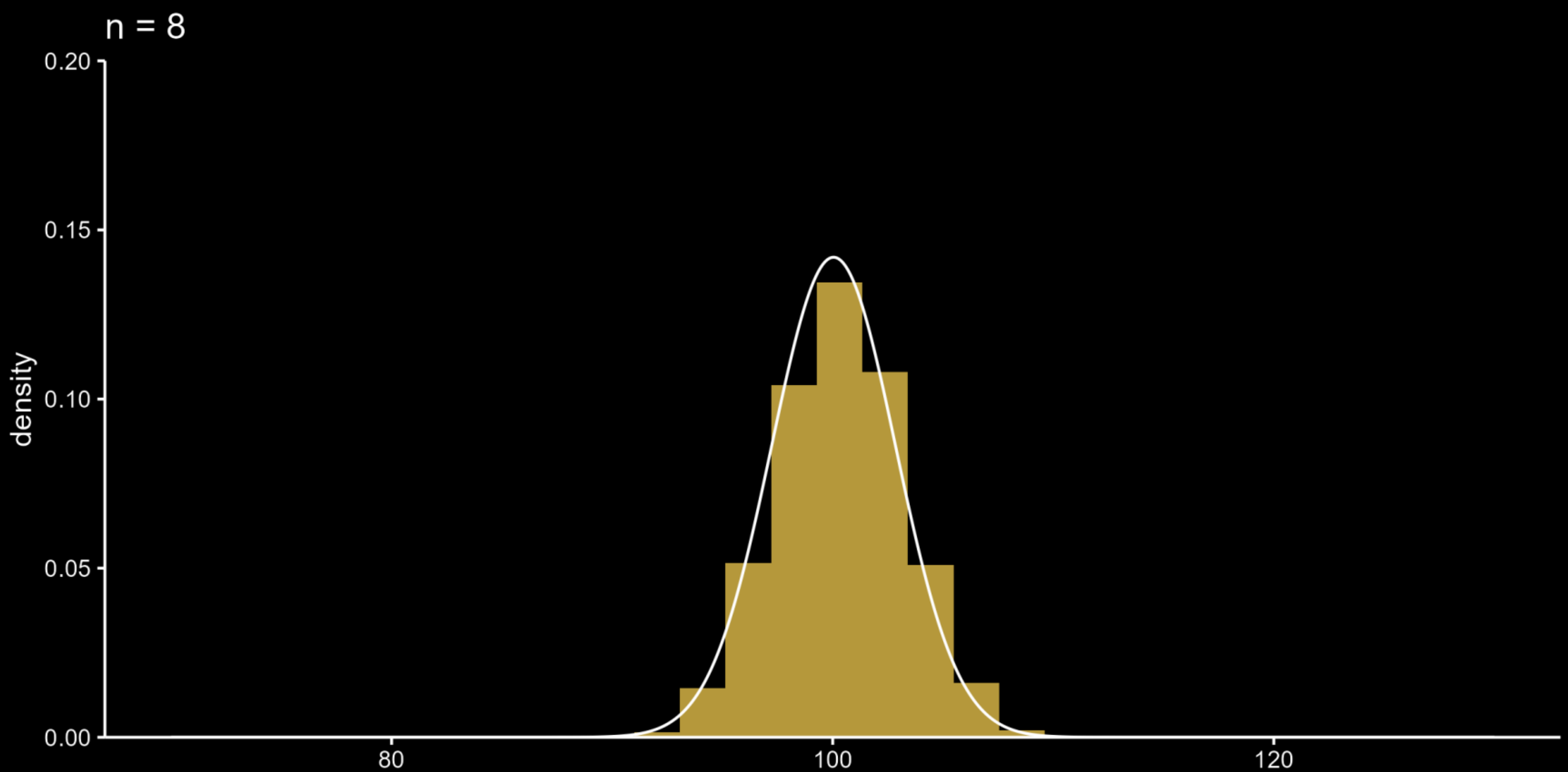

n = 5
0.20
0.15
0.10
0.05
0.00
density
80
100
120



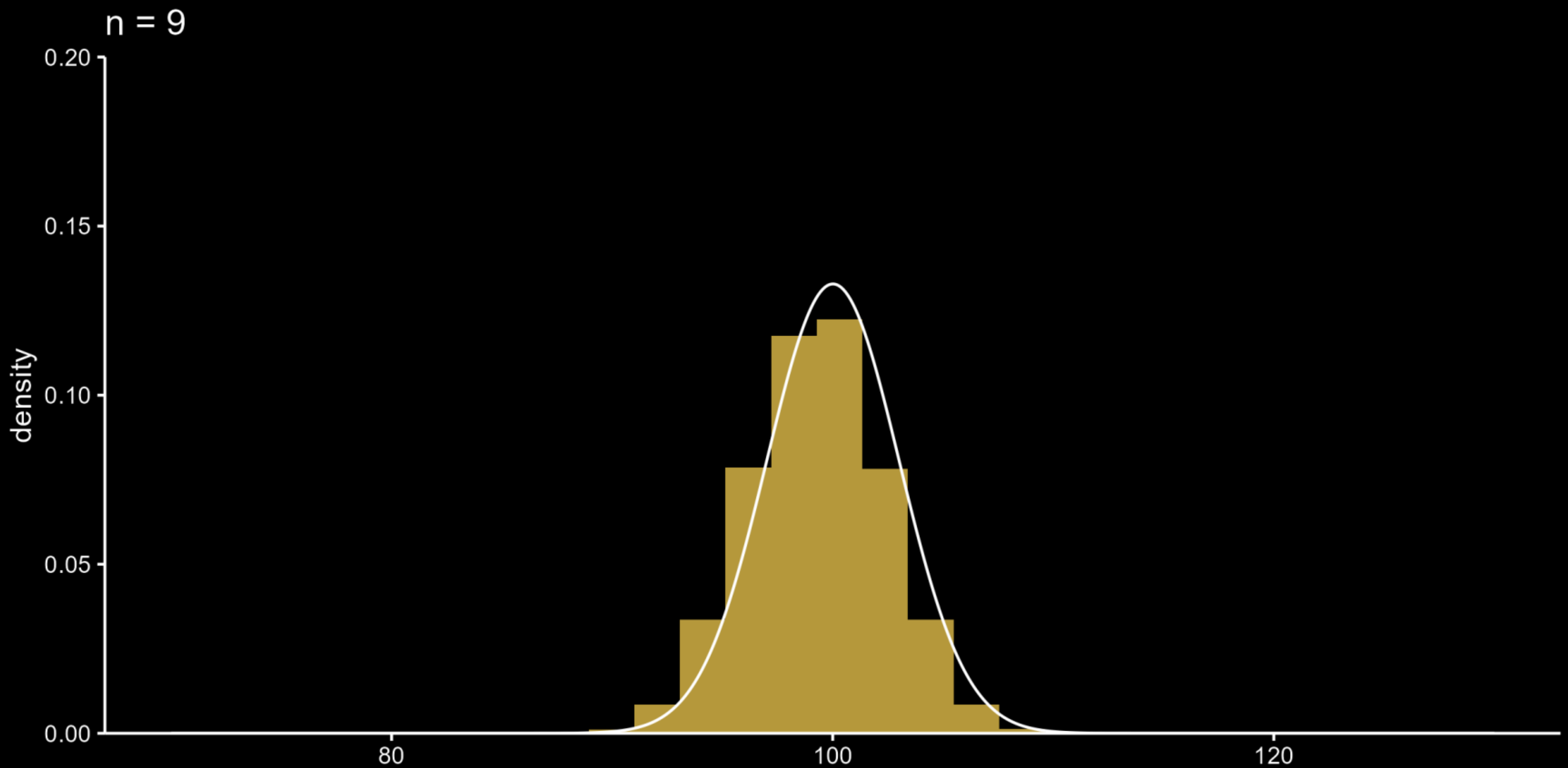
n = 6
0.20
0.15
0.10
0.05
0.00
density
80
100
120

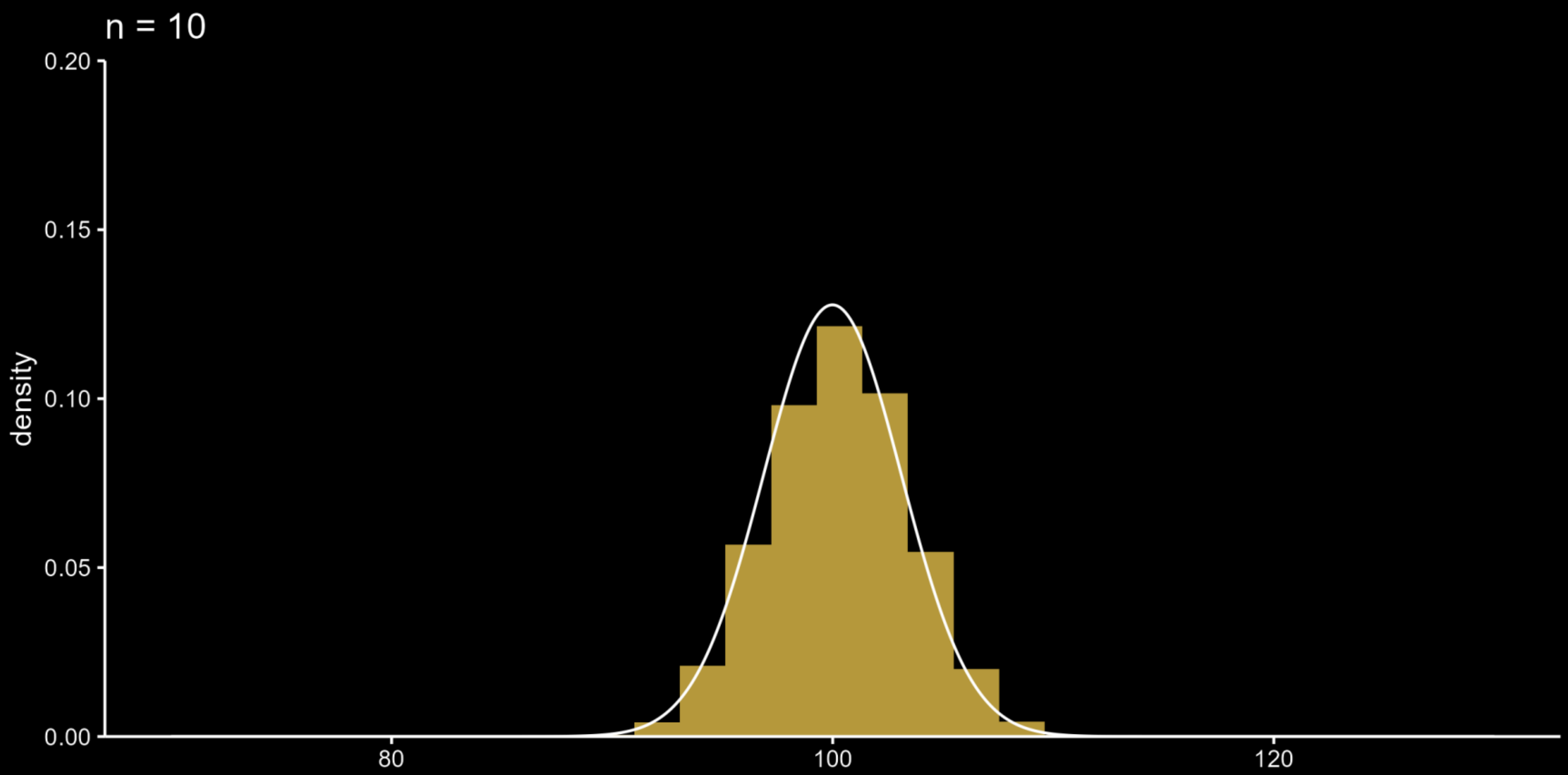


NHH
TECH3

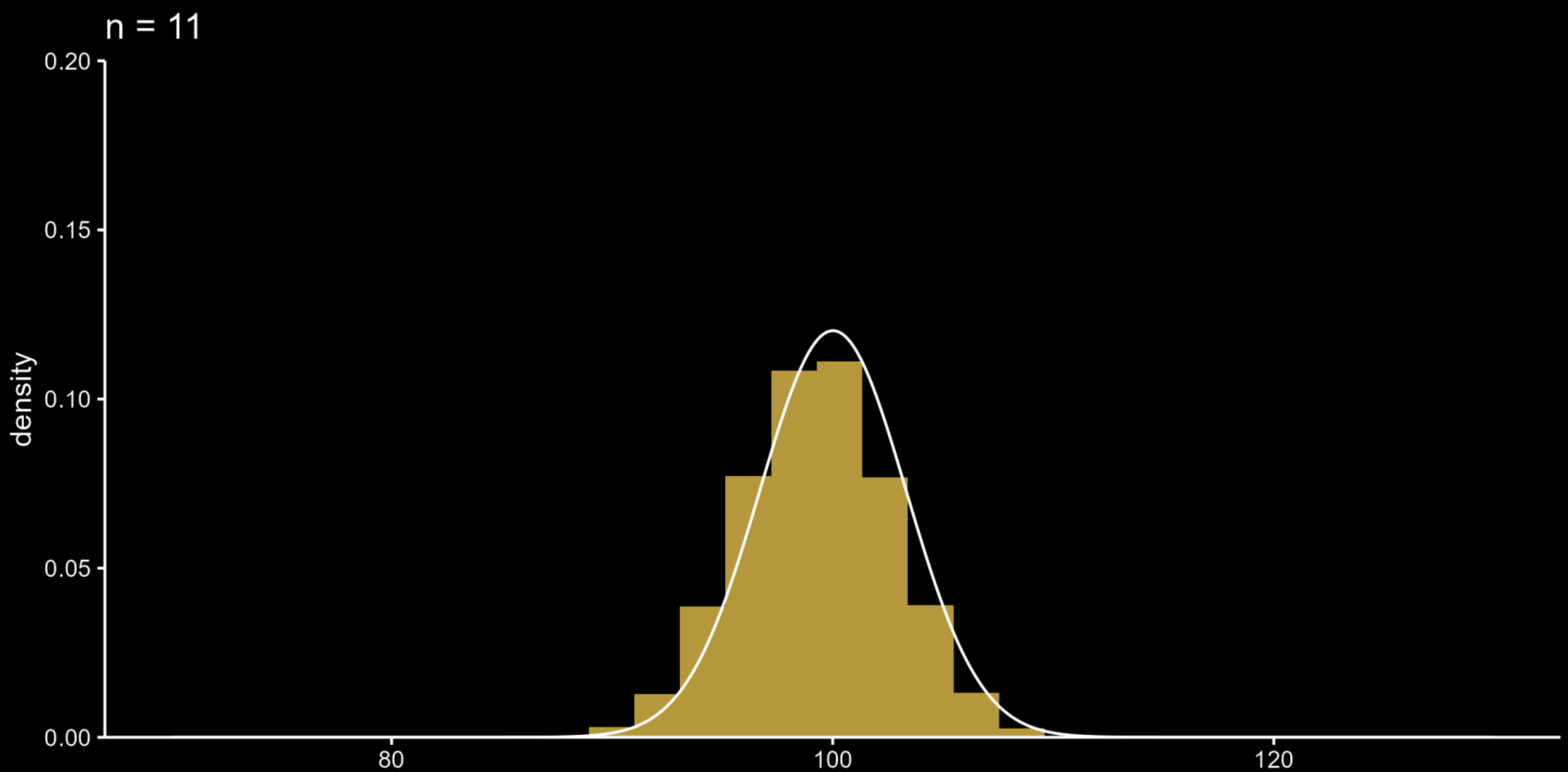


n = 8
density
0.20
0.15
0.10
0.05
0.00
80
100
120

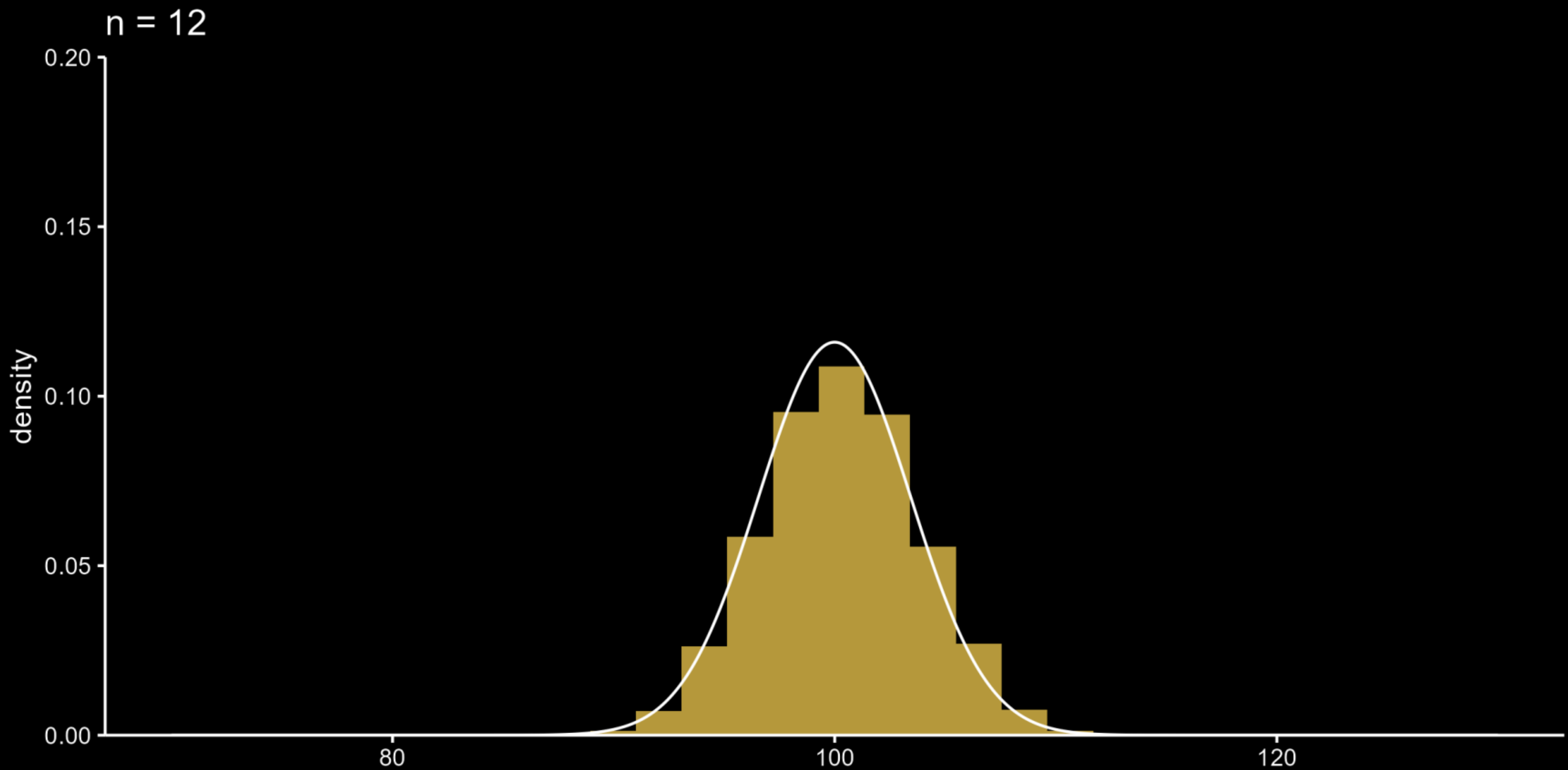

n = 9
0.20
0.15
0.10
0.05
0.00
density
80
100
120



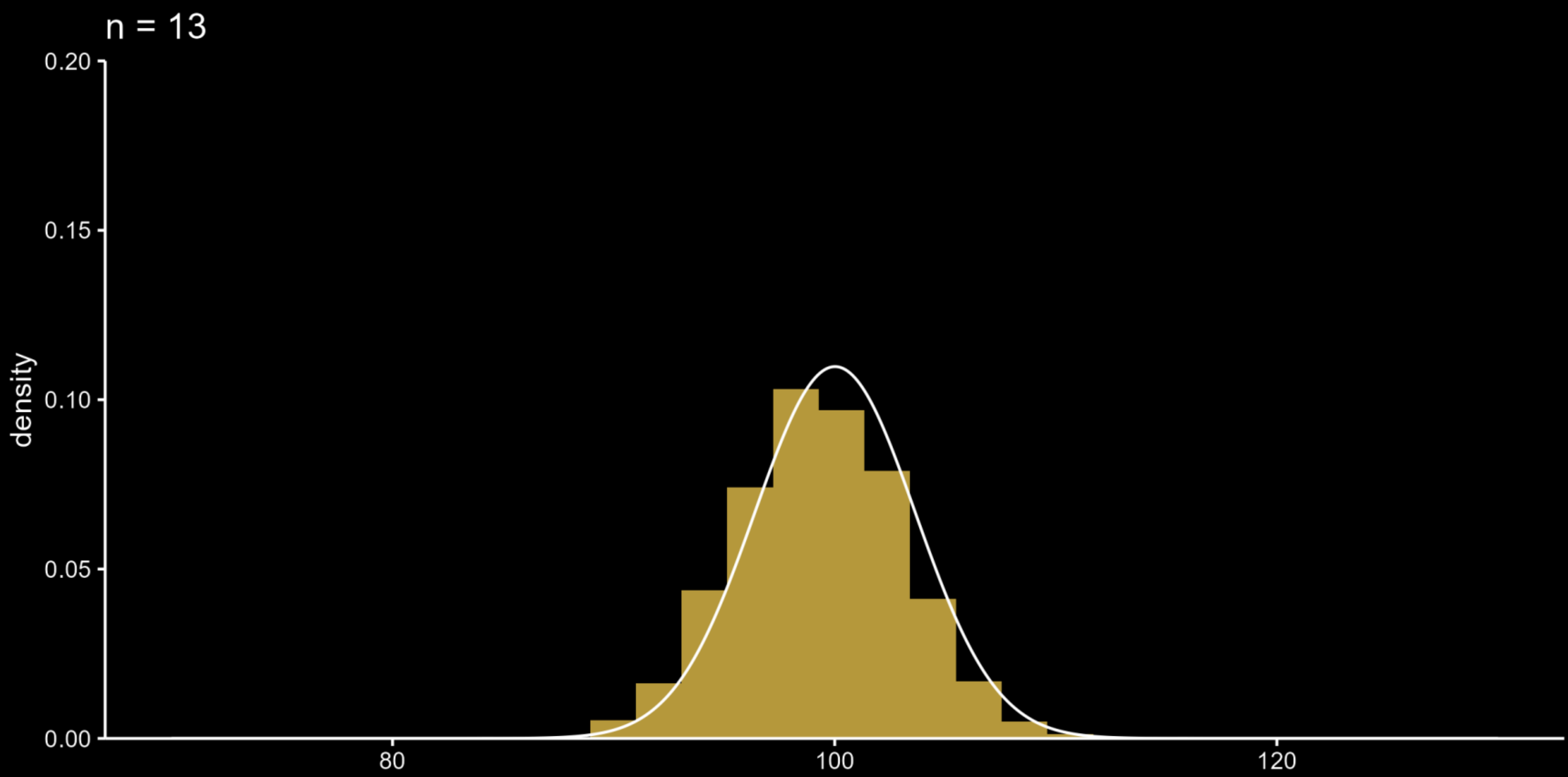
NHH
TECH3



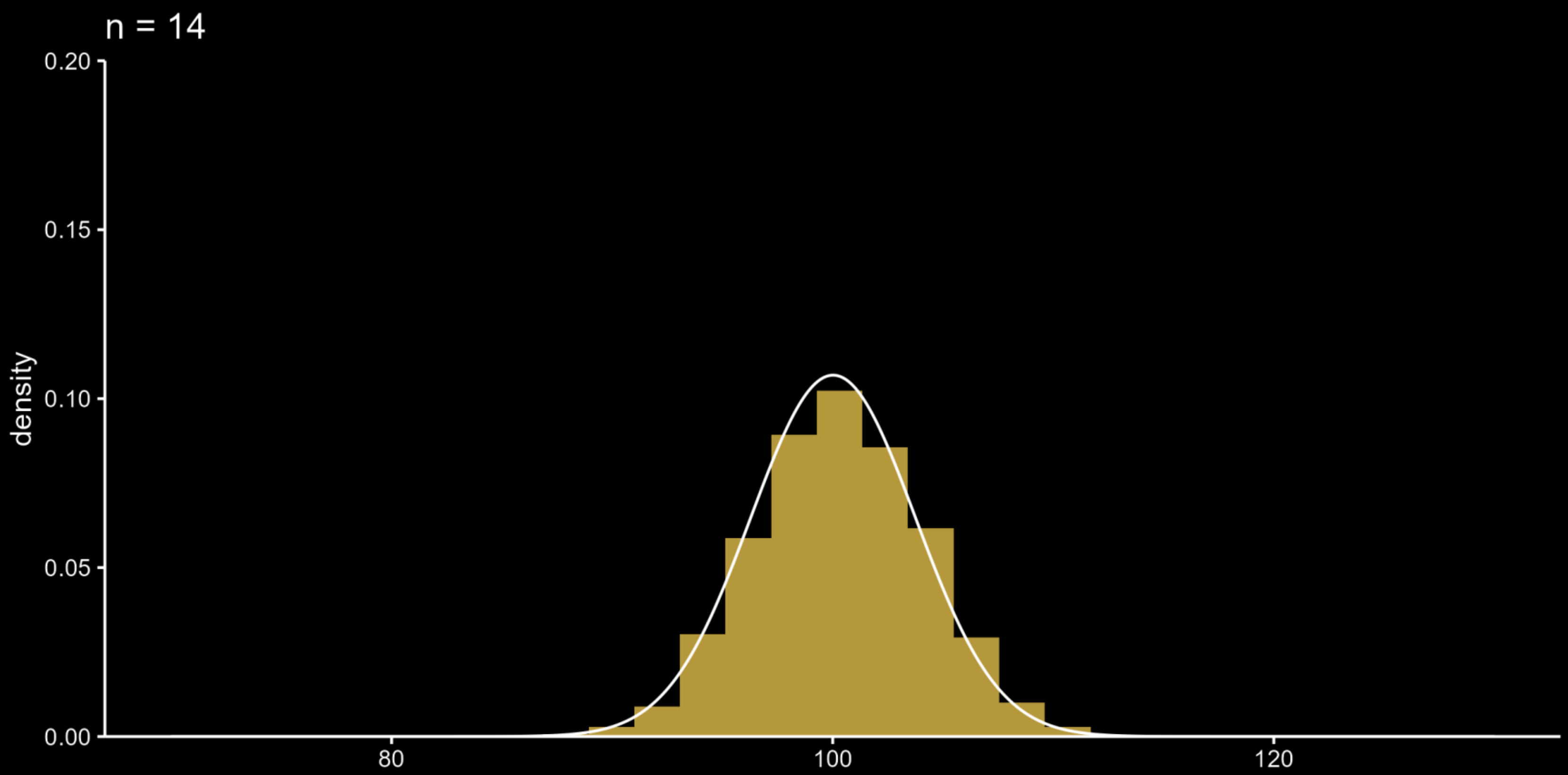
n = 11
density
0.20
0.15
0.10
0.05
0.00
80
100
120

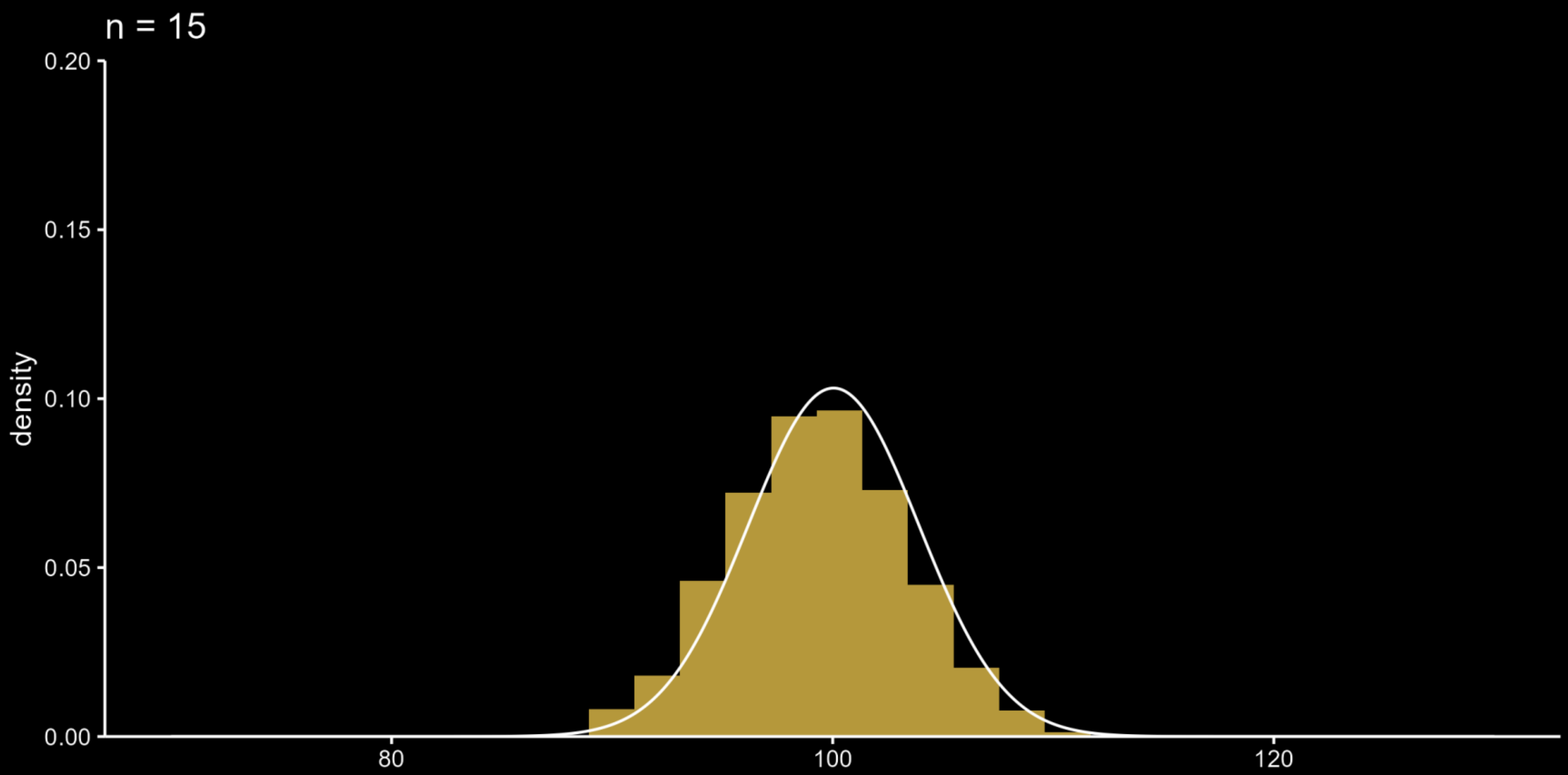


NHH
TECH3

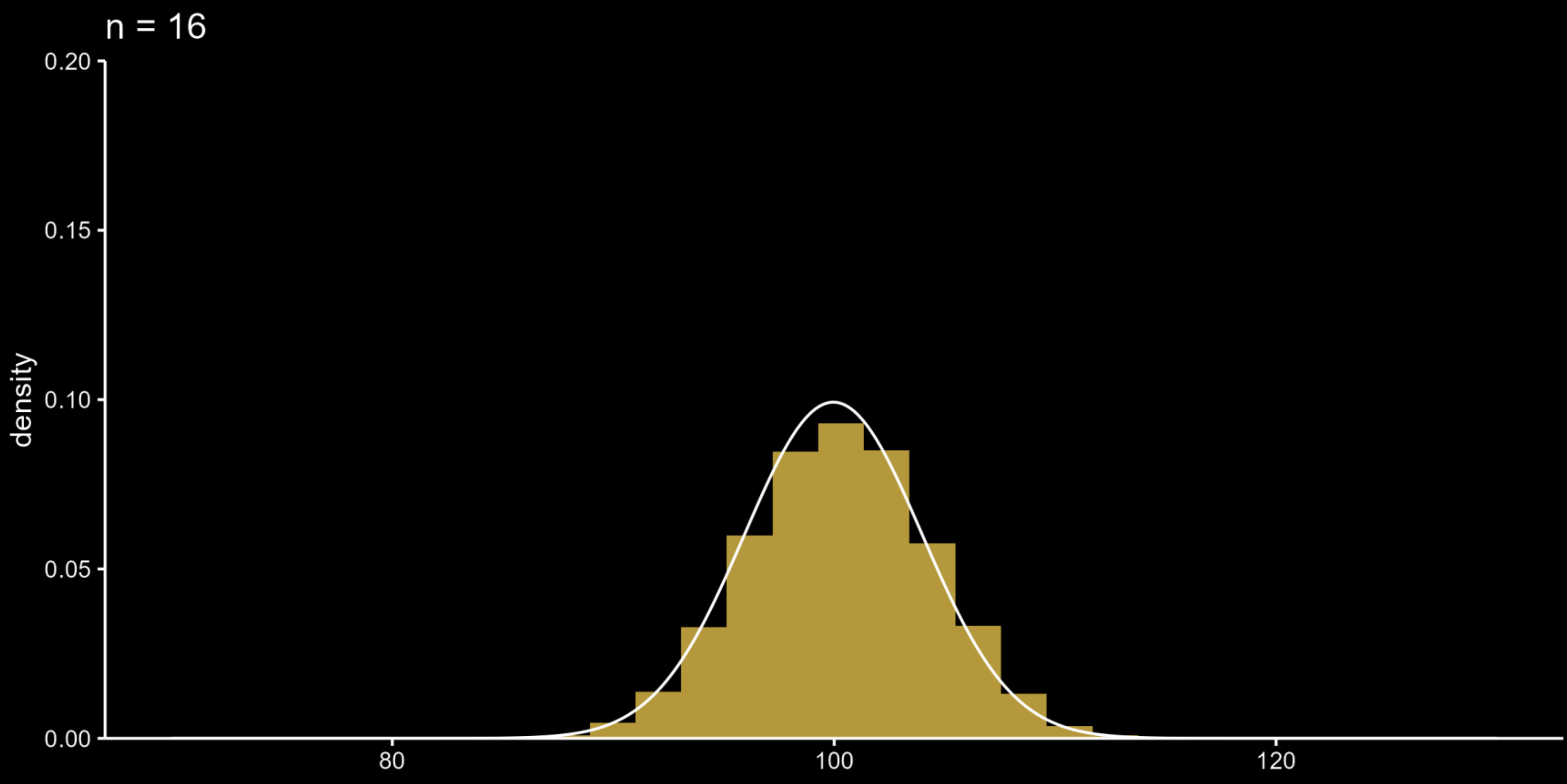


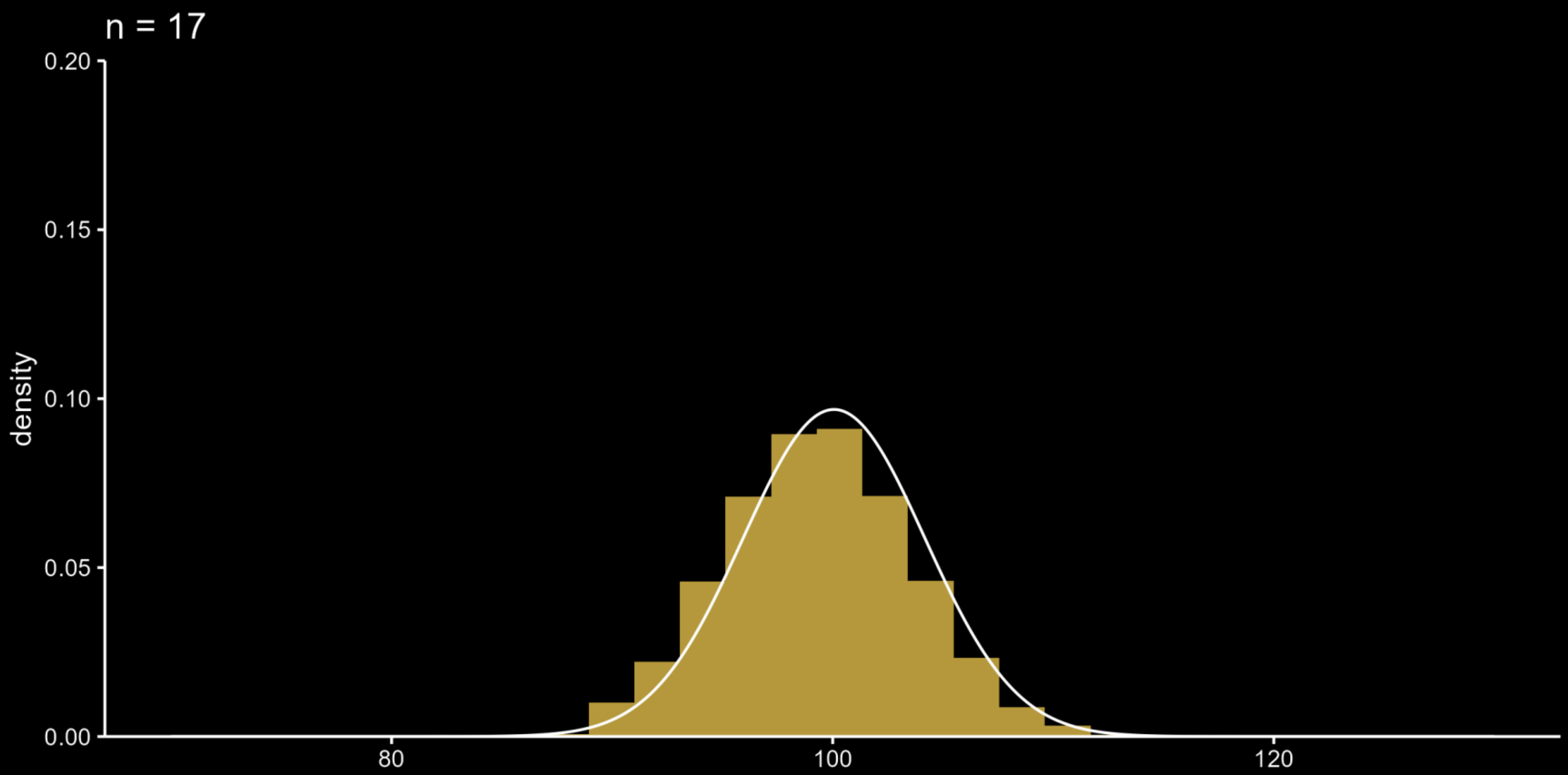
n = 13
0.20
0.15
0.10
0.05
0.00
density
80
100
120


n = 14
density
0.20
0.15
0.10
0.05
0.00
80
100
120

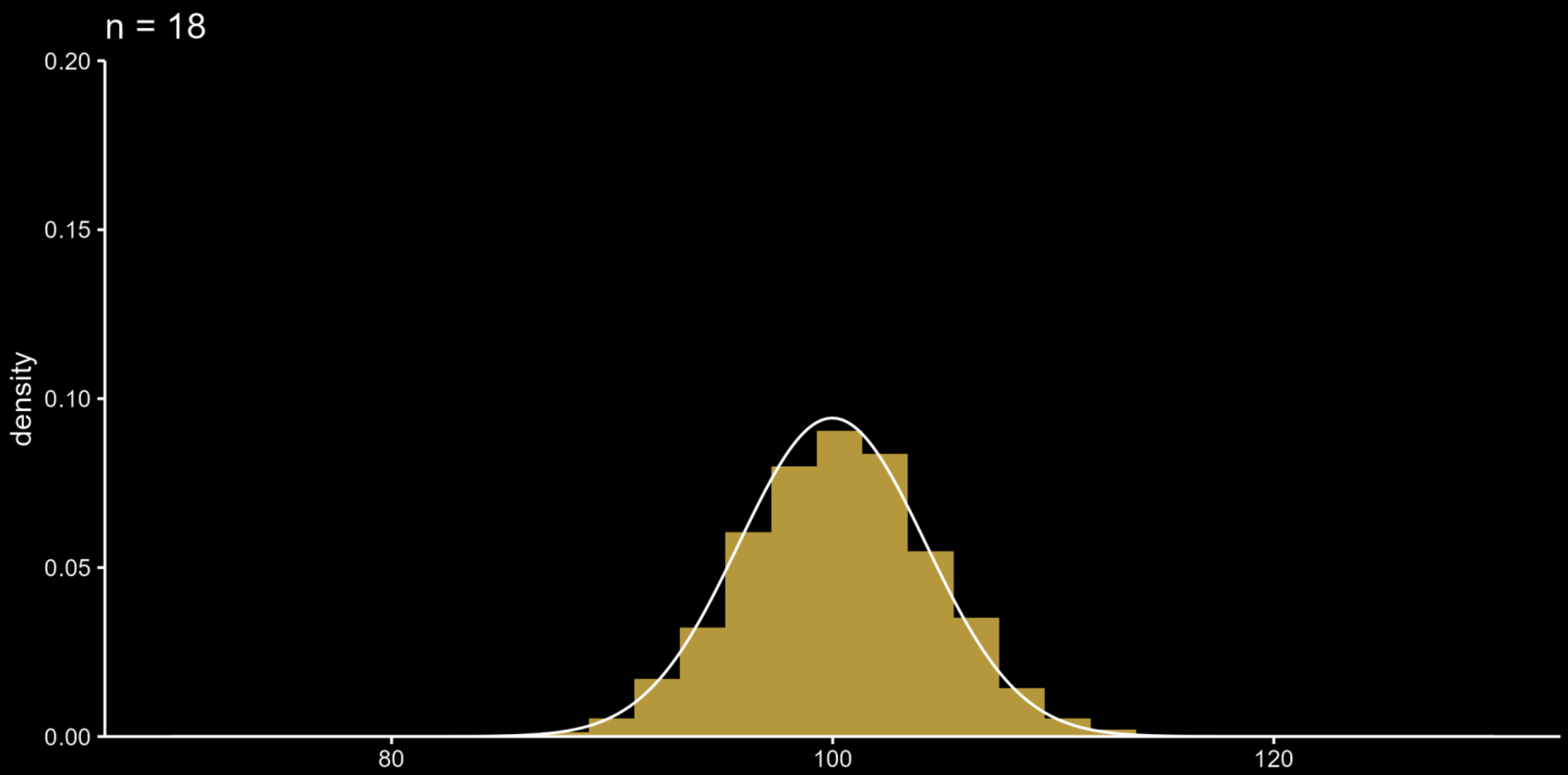


NHH
TECH3

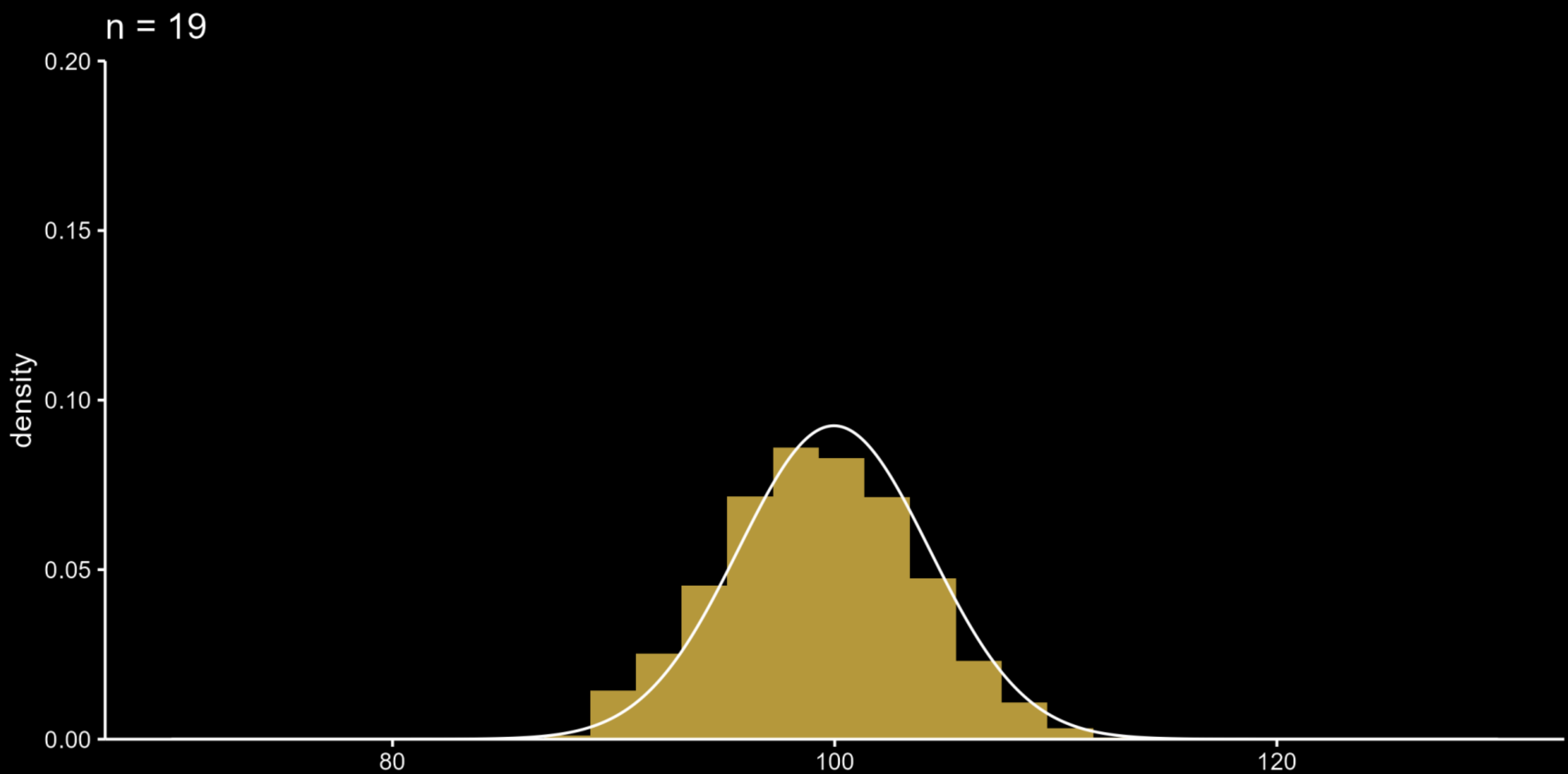

n = 16
density
0.20
0.15
0.10
0.05
0.00
80
100
120



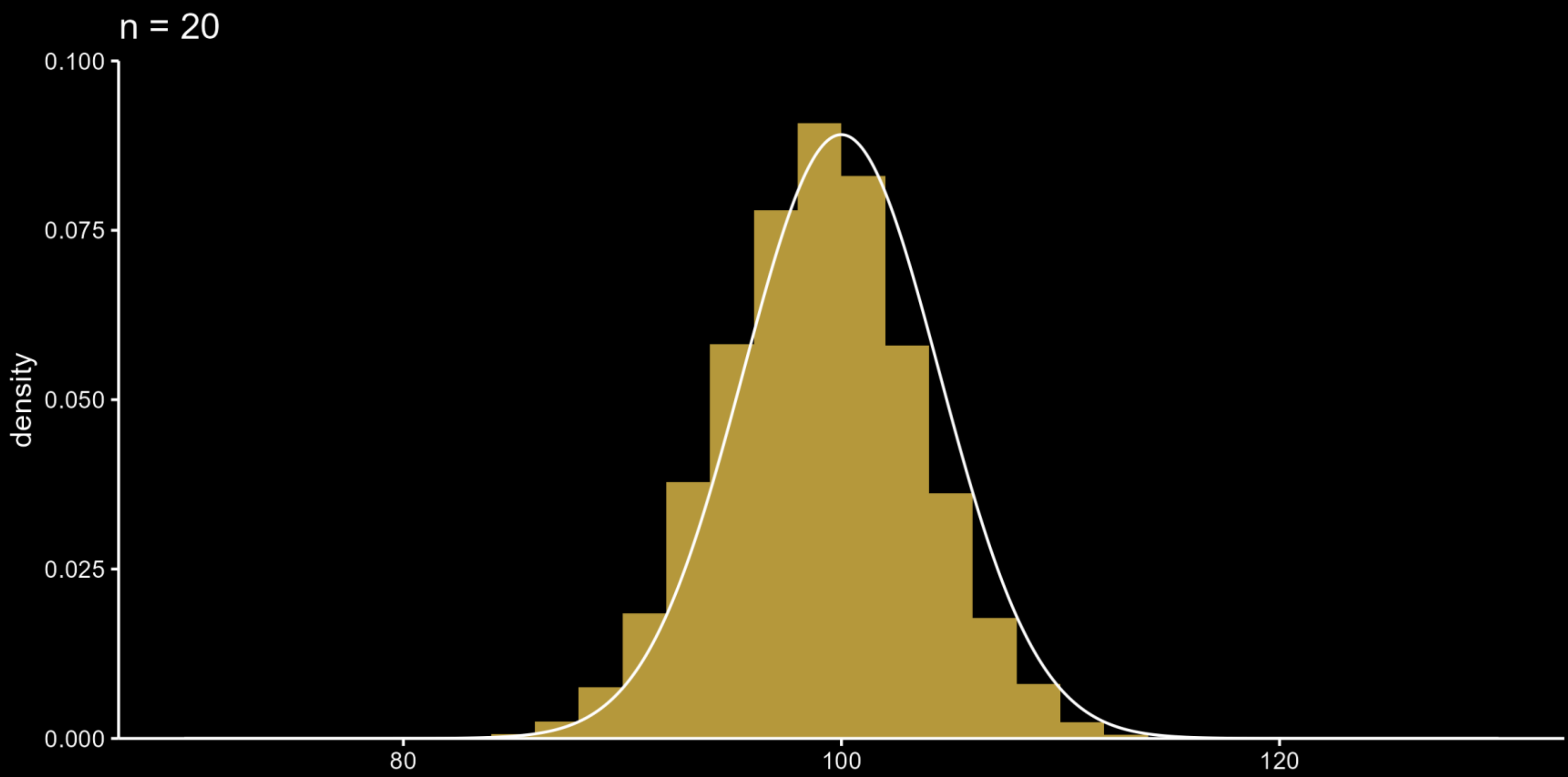
NHH
TECH3



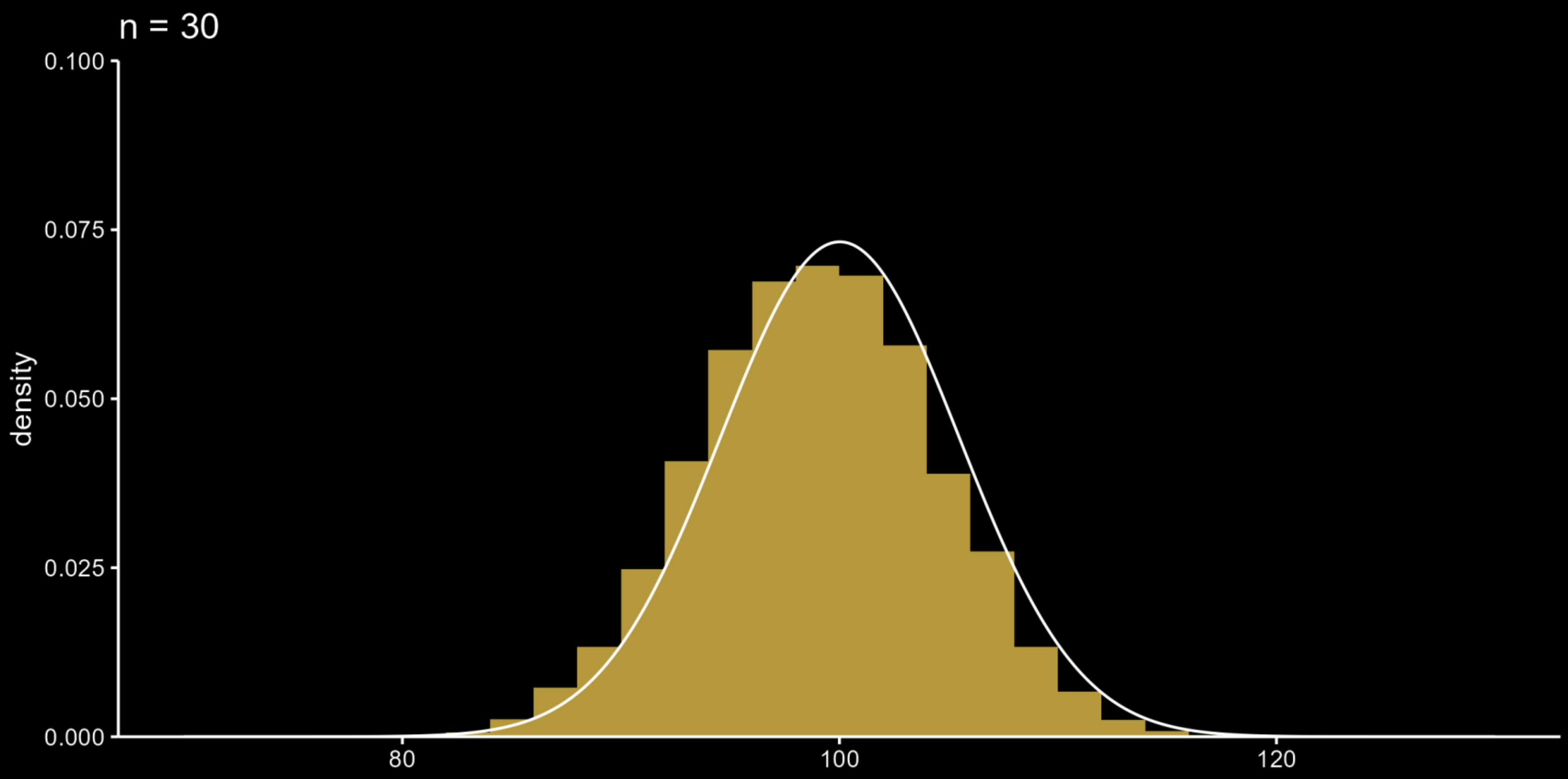
NHH TECH3



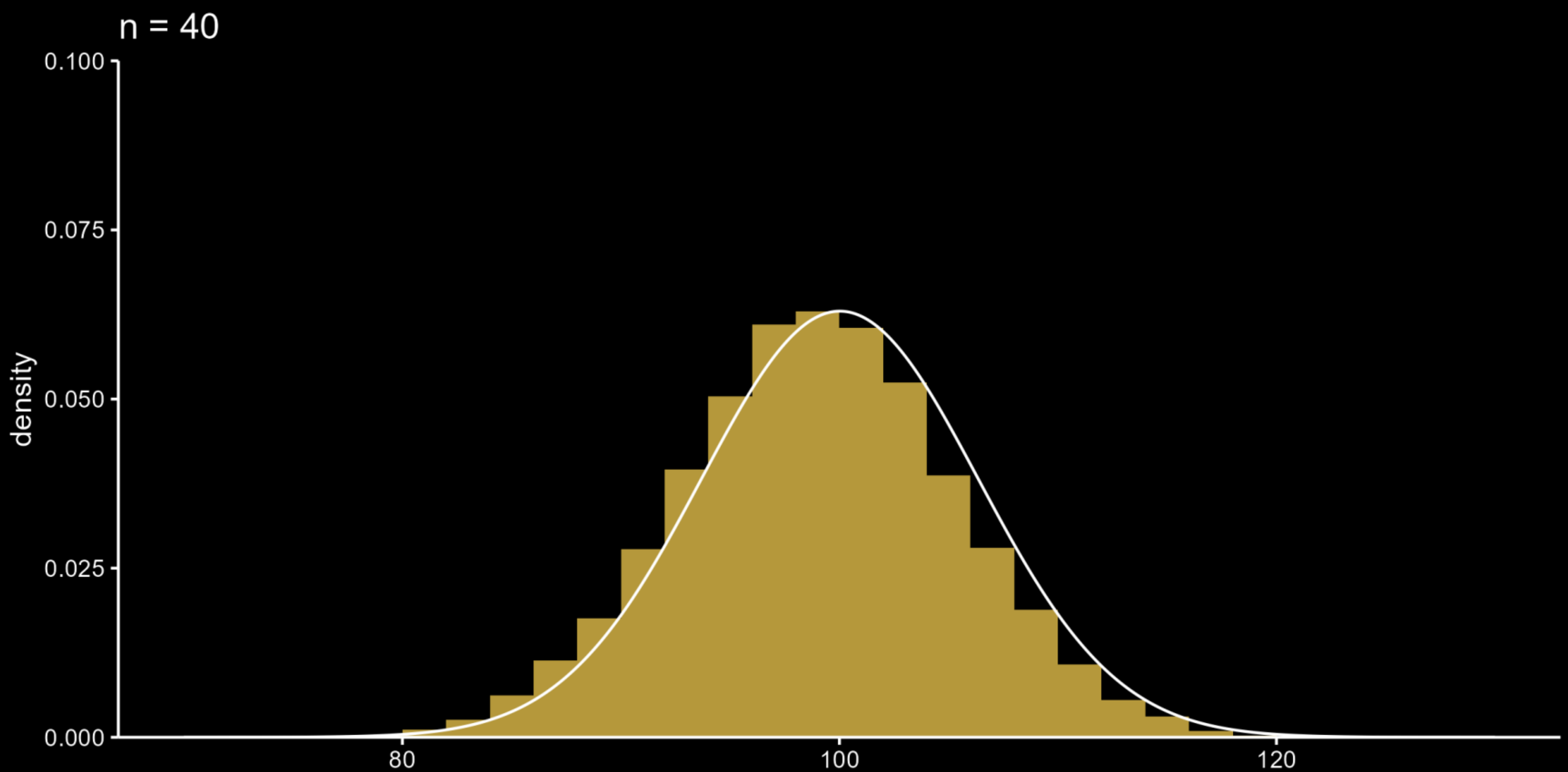
NHH
TECH3



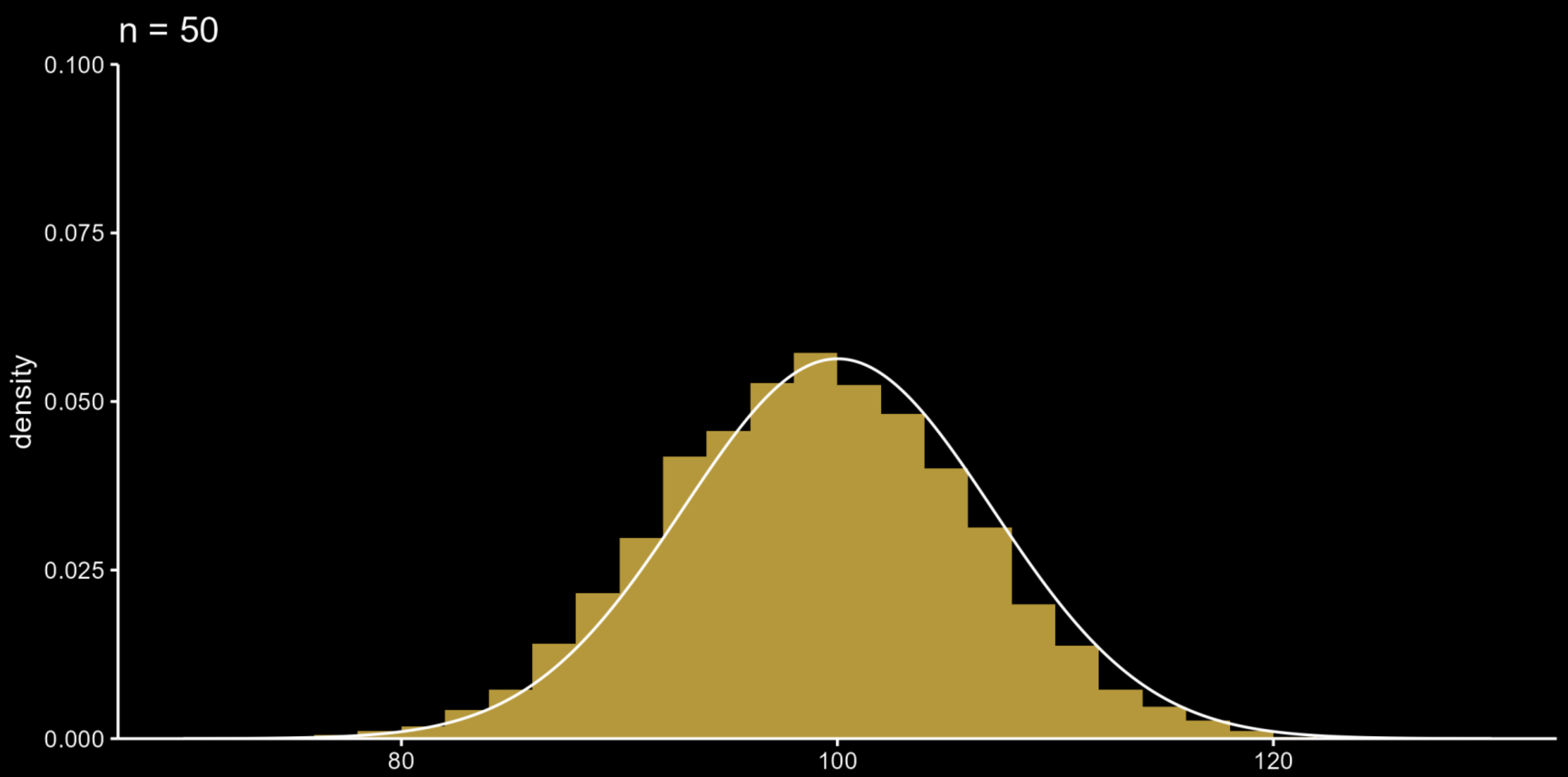
n = 20
0.100
0.075
0.050
0.025
0.000
density
80
100
120



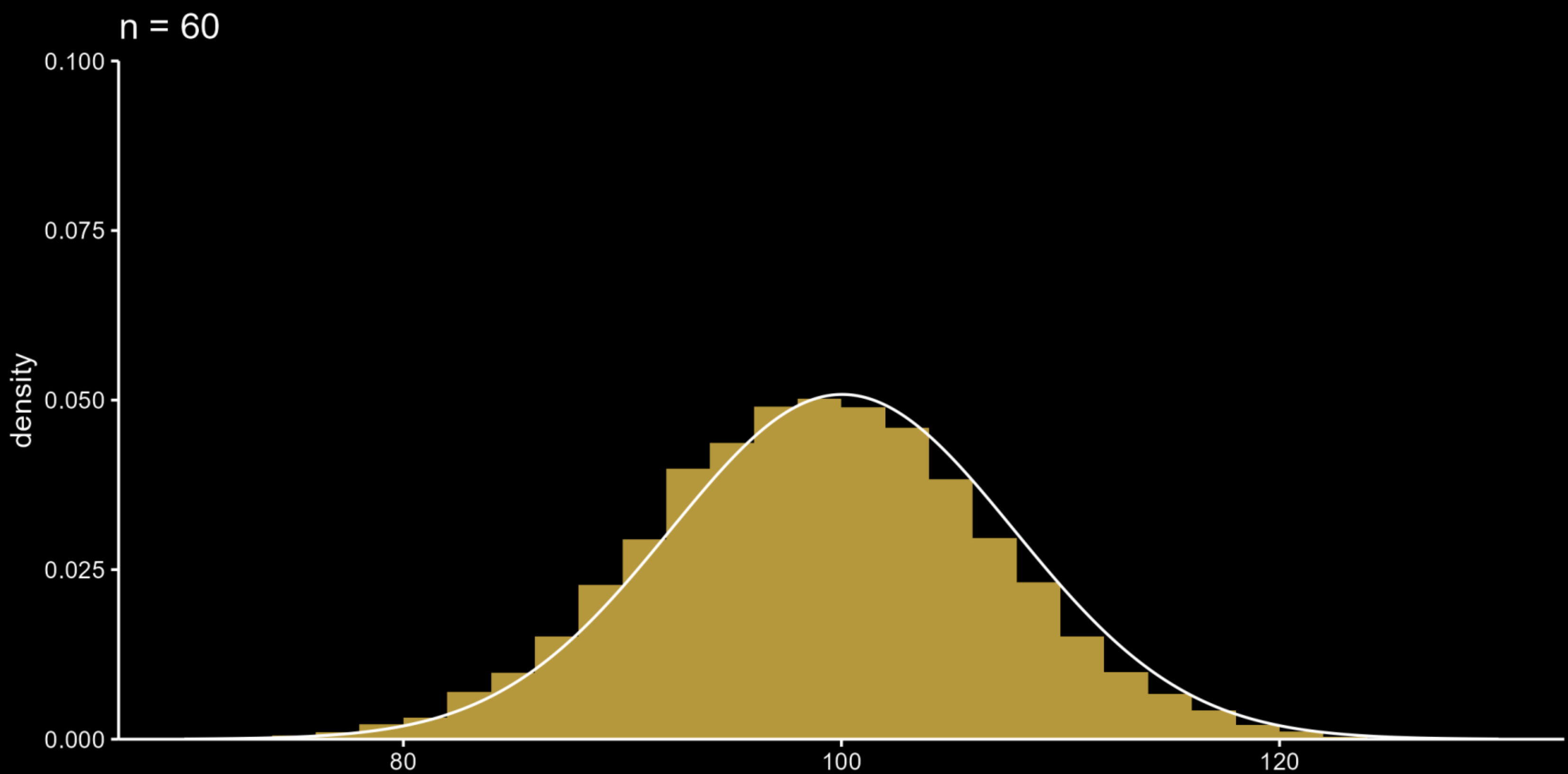
NHH
TECH3



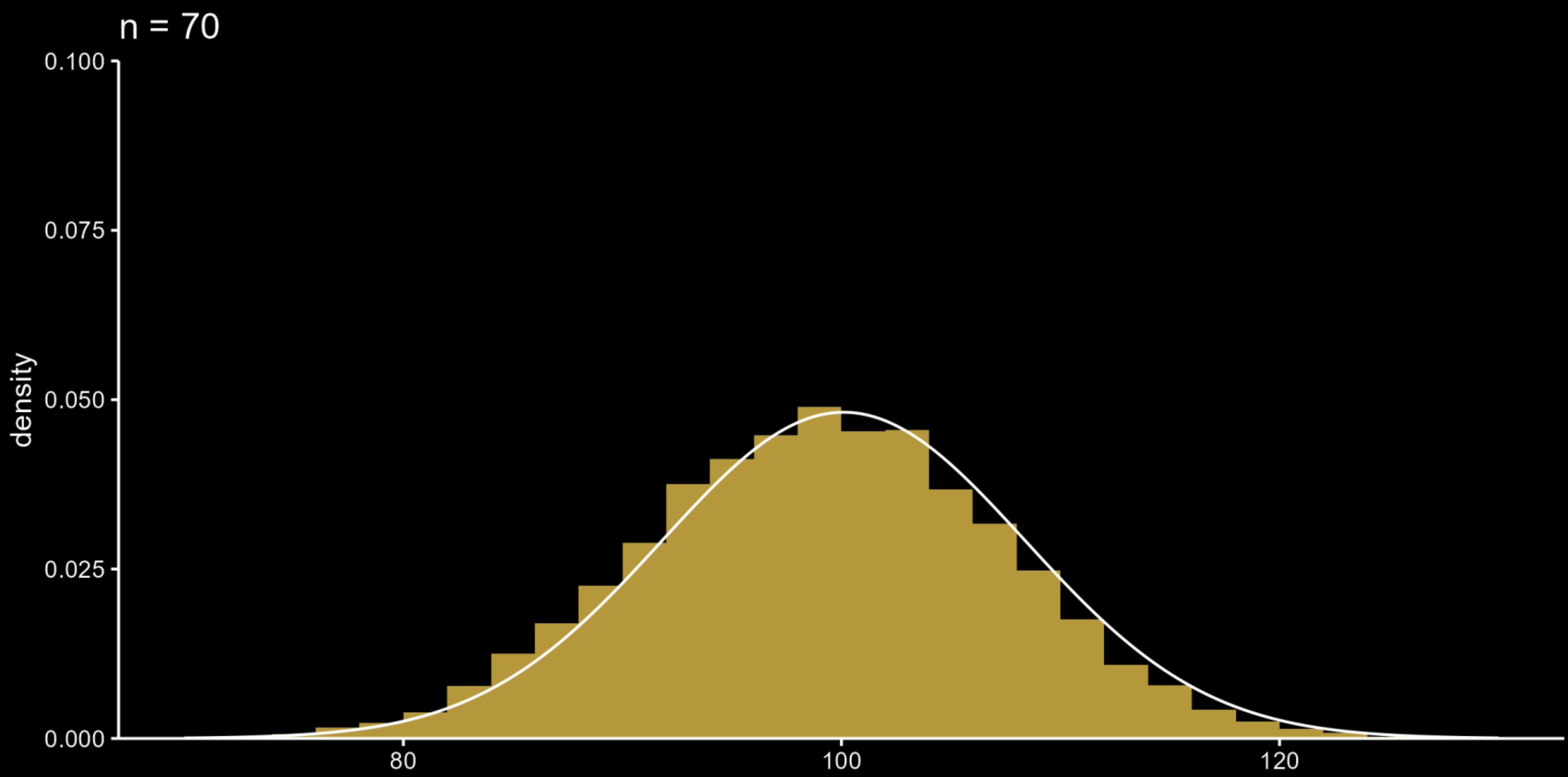
NHH TECH3



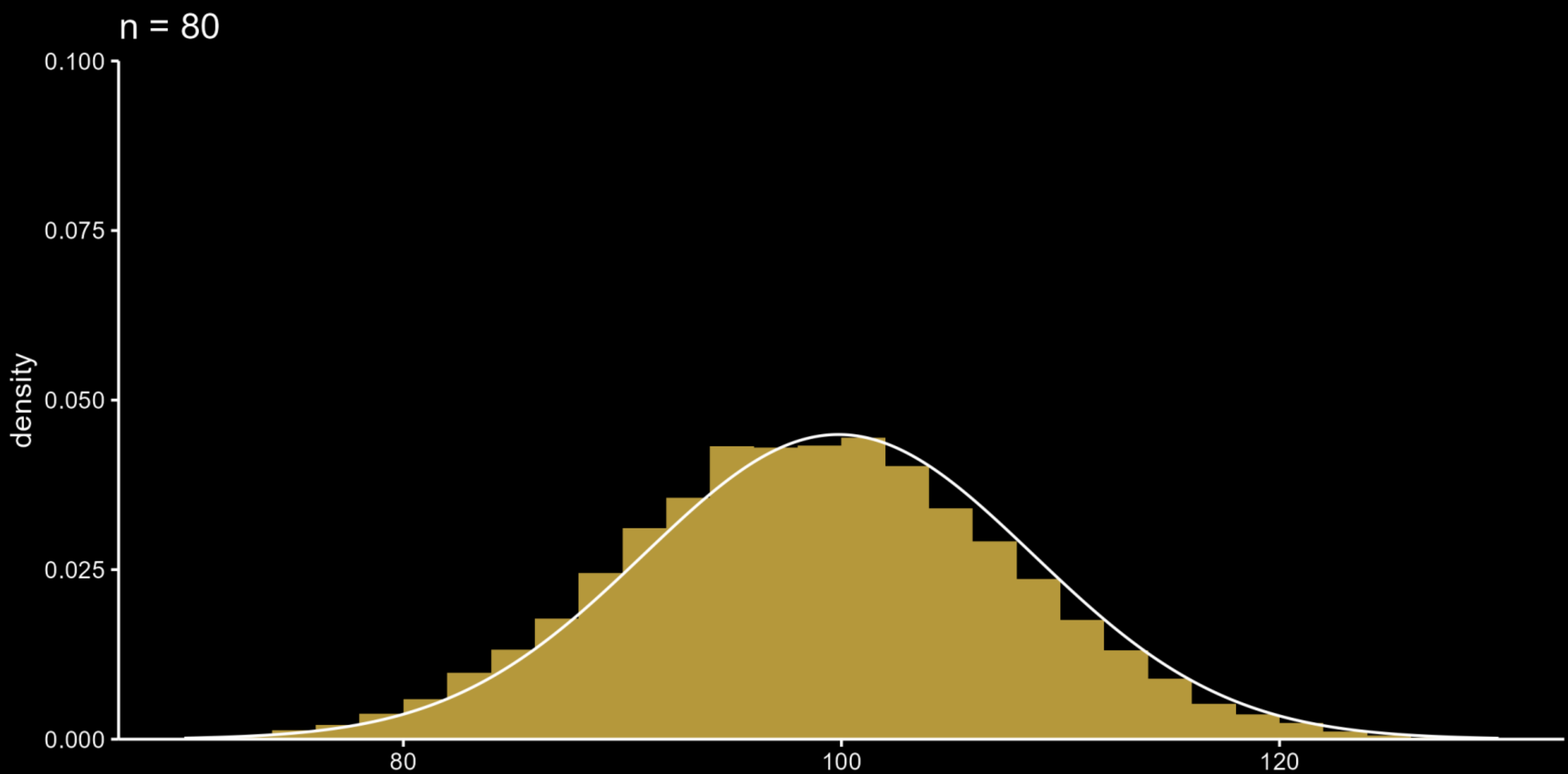
n = 50
density
0.100
0.075
0.050
0.025
0.000
80
100
120

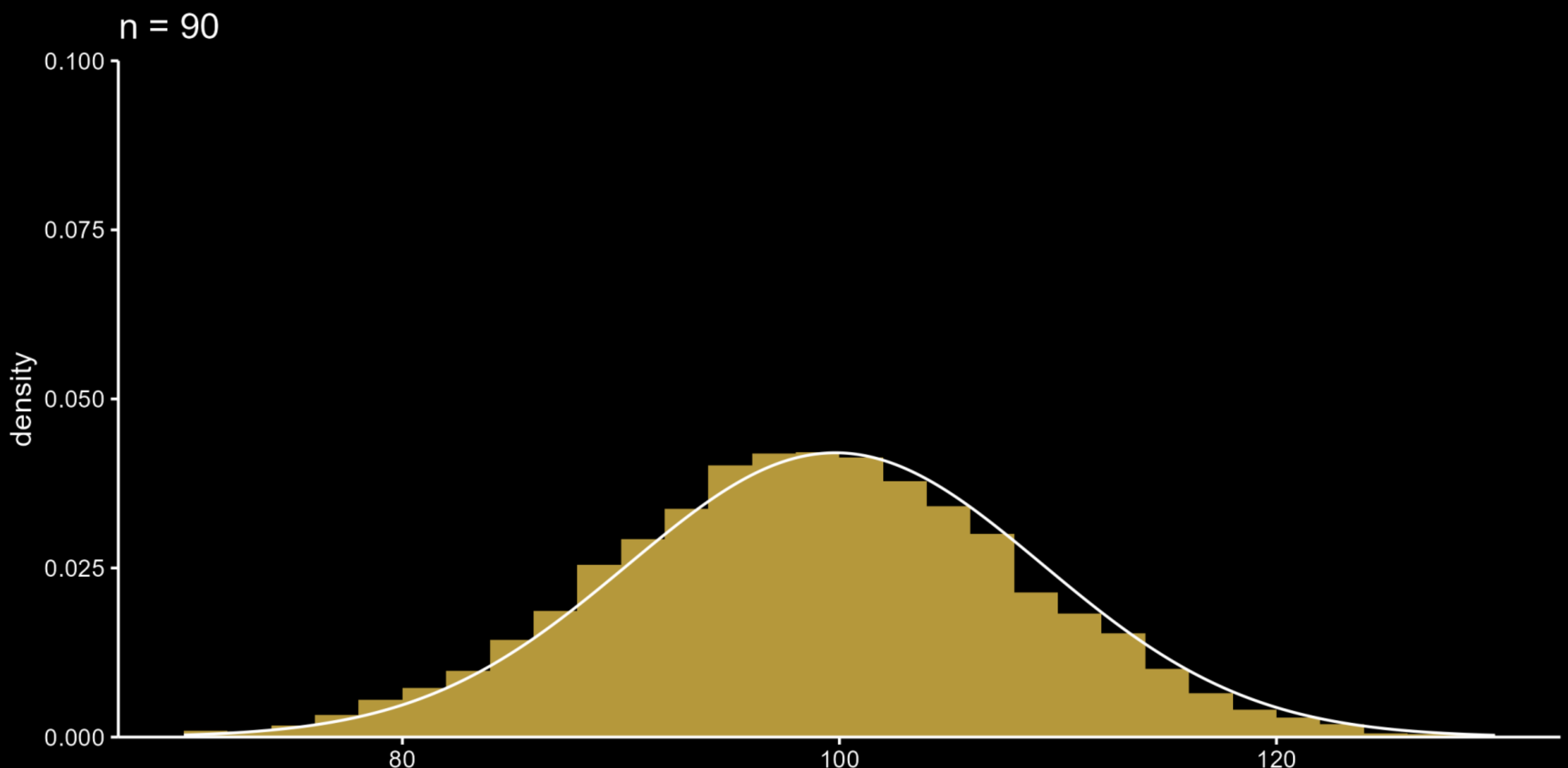


NHH
TECH3

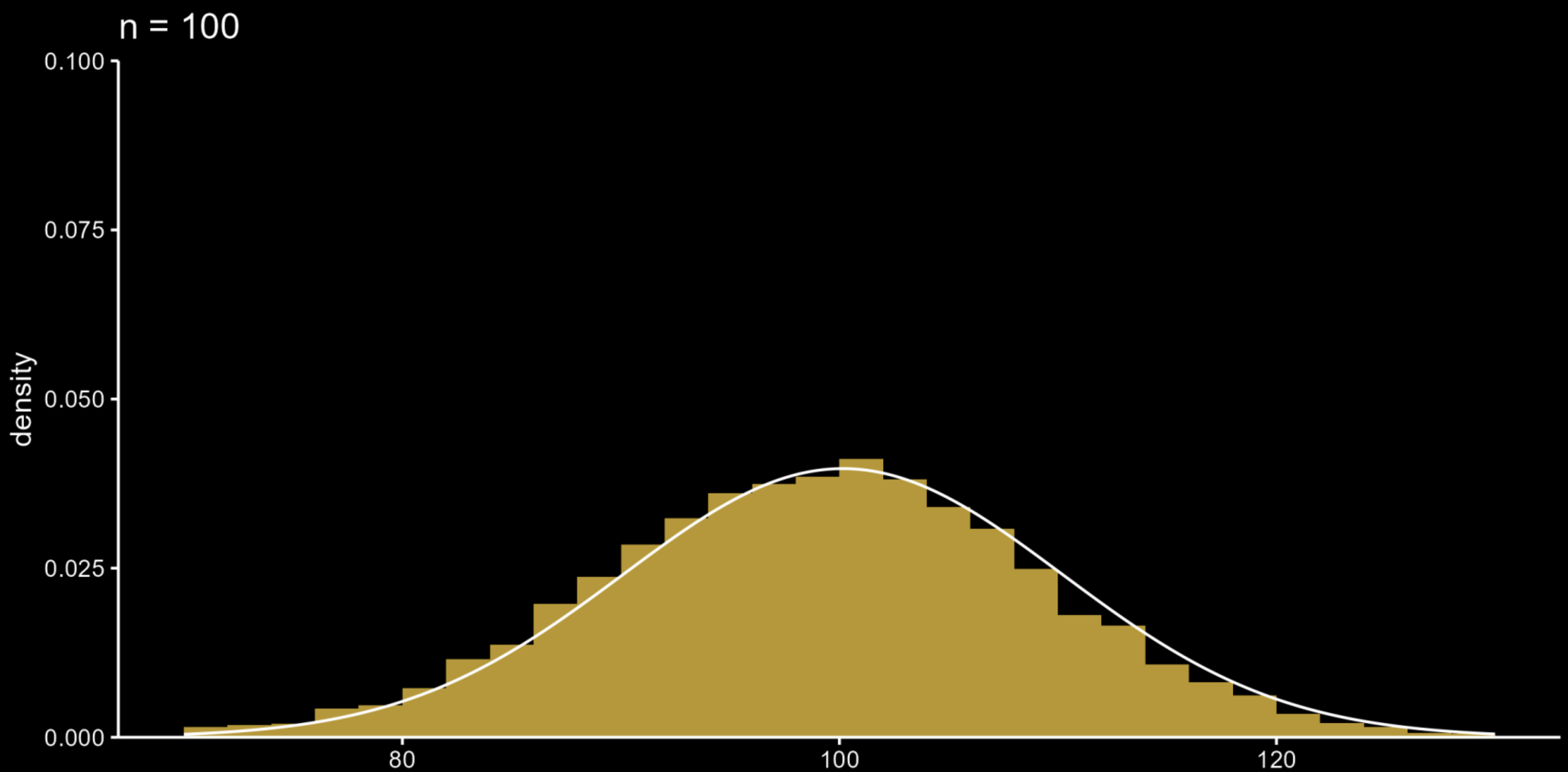


NHH
TECH3


n = 80
0.100
0.075
0.050
0.025
0.000
density
80
100
120



NHH
TECH3



NHH
TECH3

NHH

TECH3

Sondre Hølleland
Geir Drage Berentsen