University of Nottingham
UK | CHINA | MALAYSIA

Introductory Econometrics BUSI2053

3. Statistics for Econometrics

Probability & Statistics

Probability: Given the information in the bucket, what is in your hand?

Statistics: Given the information in your hand, what is in the bucket?

Image by MIT OpenCourseWare. Based on Gilbert, Norma. Statistics. Philadelphia, PA: W. B. Saunders Co., 1976.

Lecture Outline

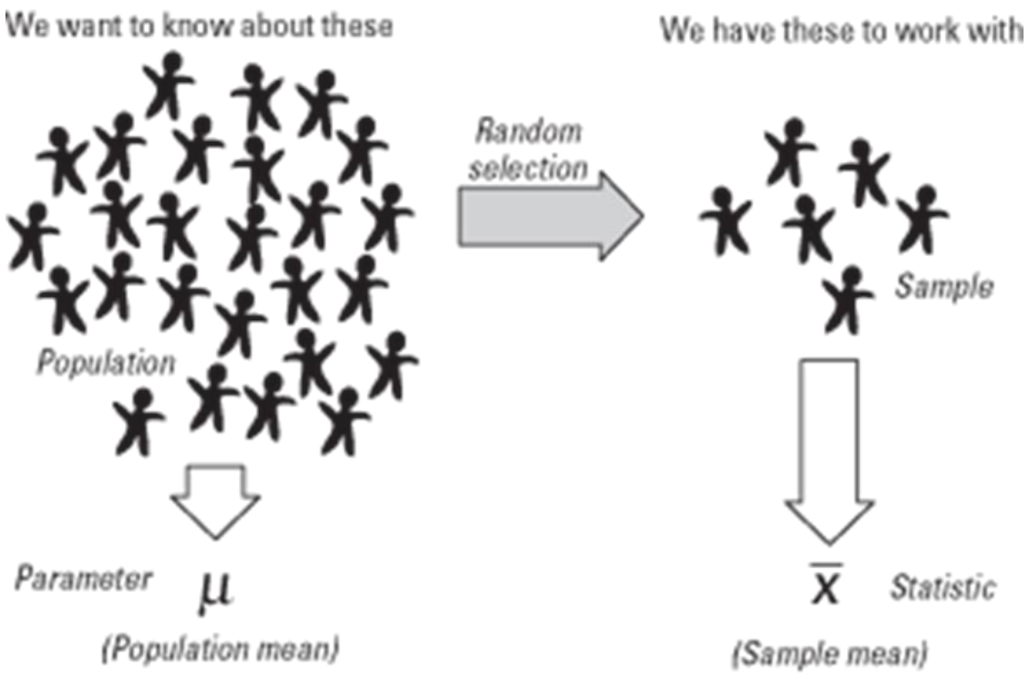
- Random variable and population
- Statistical Inference
- Sampling Distribution
- Estimate and estimator
- Expected value of sample means
- Variance of sample means
- Properties of an Estimator

- **Suggested Reading:**
- Appendix C: Hill, R.C., Griffiths W.E. and Lim, G.C. Principles of Econometrics, fourth edition, Wiley, 2012 (pp. 692-716)
- Appendix A: Gujarati, D.N. and Porter D.C. Basic econometrics, 5th ed., McGraw-Hill, 2009 (pp. 801-823)
- Review: Dougherty, Christopher. Introduction to econometrics. 4th ed. Oxford University Press, 2011 (pp.5-72)
- Chapter 3.1, 3.2, 4.1: Westhoff, F. An introduction to econometrics: a self-contained approach, MIT Press, 2013

Fundamentals of Statistics

Statistics transform numbers into useful information that we can use in decision making.

Population: A population is any specific collection of objects of interest.
Sample: Any subset or sub-collection of the population selected for analysis.
Variable**:** A characteristic of an item or an individual that will be analysed using statistics.
Parameter: A numerical measure that describes a variable (characteristic) of a population.
Statistic: A numerical measure that describes a variable (characteristic) of a sample



https://forms.office.com/r/PpuJW9G9SM

Random variable and population

- We assume that the population has a center, which we describe by the expected value of the random variable X:

$$E(X) = \mu$$

  – μ is a **population parameter**, or, more briefly, a parameter

- The other random variable characteristic of interest is its variability, which we measure by its variance:

$$\mathrm{var}(X) = \mathrm{E}[X - E(X)]^2 = E[X - \mu]^2 = \sigma^2$$

- Denote the mean and variance of a random variable as:

$$X \sim (\mu, \sigma^2)$$

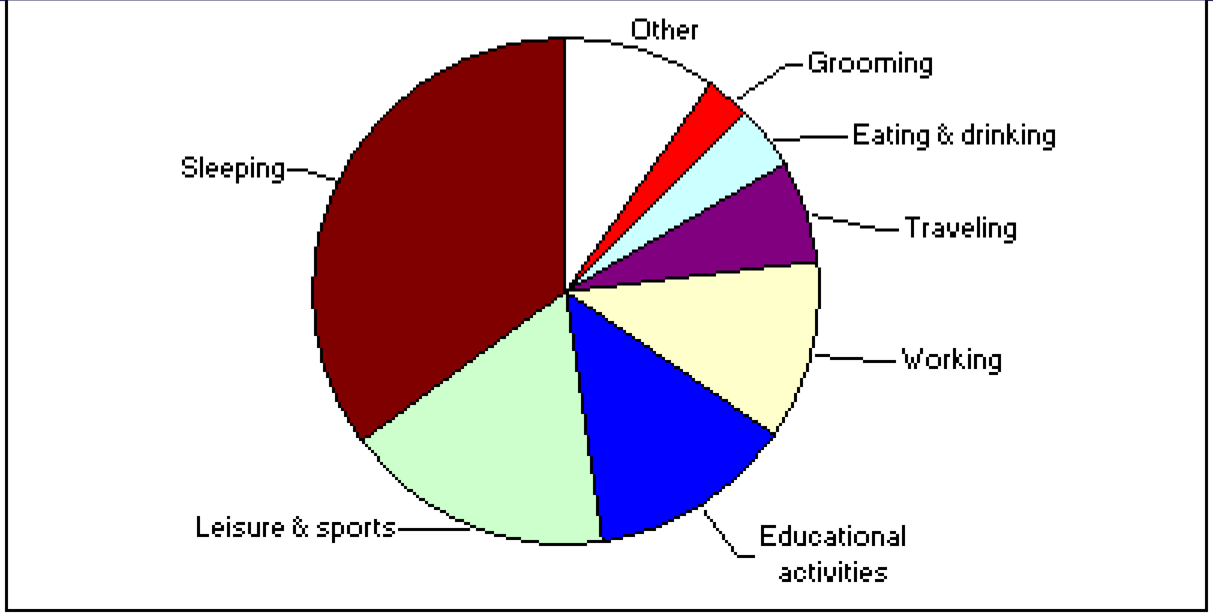
Random variable and population

- If we select individual observation from a population randomly, we can assume that $x_1, x_2, \ldots, x_N$ are statistically independent with identical probability distributions
- It is sometimes reasonable to assume that population values are normally distributed, which we represent by:

$$X \sim N(\mu, \sigma^2)$$

Statistical inference

- Decision makers are interested in economic variables
- Sometimes questions of interest focus on a single economic variable
  - e.g., Policy makers would be interested in time usages of students
- But we cannot ask each and every student
- This is a situation when statistical inference is used
- Draw conclusions about a population based on a sample of data



Full-time university students spent (on average) on a weekday:

- 3.1 hours engaged in educational activities
- 8.5 hours sleeping
- 4.1 hours in leisure and sports activities
- 2.7 hours working
- 1.5 hours for traveling,
- 1.0 hour for eating and drinking
- 0.7 hour for grooming

Source: Bureau of Labor Statistics, U.S. Department of Labor, The Economics Daily, College students and time use, 2003-2005 at https://www.bls.gov/opub/ted/2007/may/wk4/art04.htm

An Estimate and Estimator

- We can use the average value in the sample, or sample mean, to estimate the population mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- The sample mean ($\bar{x}$) is **an estimate** of the population mean (which is unknown)
- The formula to calculate the sample mean is called **an estimator.** To distinguish between the estimate and the estimator of the population mean μ we will write the estimator as

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- An estimate is a **random variable** because its value is unknown until it is observed

An Estimate and Estimator

Suppose, we have recruited 4 new office staff, and their ages are as follows:

Staff	Age
1	22
2	24
3	26
4	28

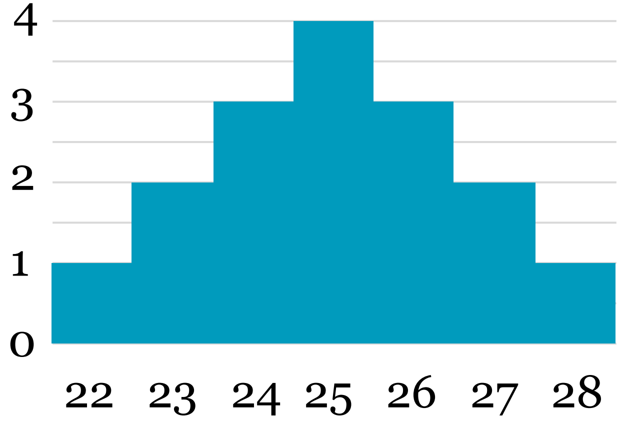
Since this is the population of our new staff, what is the population mean & variance?

$\mu = 25; \sigma^2 = 5$

From the population, we take all possible random samples of size two with replacement

Samples	Picks	$\bar{x}$
S1	22, 22	22
S2	22, 24	23
S3	22, 26	24
S4	22, 28	25
S5	24, 22	23
S6	24, 24	24
S7	24, 26	25
S8	24, 28	26
S9	26, 22	24
S10	26, 24	25
S11	26, 26	26
S12	26, 28	27
S13	28, 22	25
S14	28, 24	26
S15	28, 26	27
S16	28, 28	28

$\bar{x}$	Freq:	$P(\bar{x})$
22	1	0.0625
23	2	0.125
24	3	0.1875
25	4	0.25
26	3	0.1875
27	2	0.125
28	1	0.0625
total	16	1

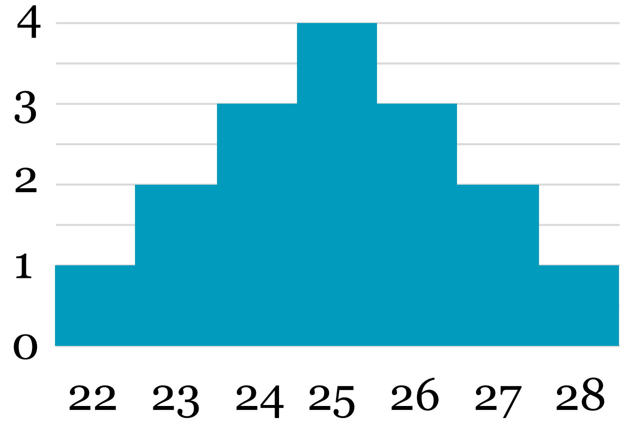


An Estimate and Estimator

- The estimates differ from sample to sample because $\bar{x}$ is a random variable
  - This variation, due to collection of different random samples, is called **sampling variation**
  - an estimator's probability density function is called its **sampling distribution**

Samples	Picks	$\bar{x}$
S1	22, 22	22
S2	22, 24	23
S3	22, 26	24
S4	22, 28	25
S5	24, 22	23
S6	24, 24	24
S7	24, 26	25
S8	24, 28	26
S9	26, 22	24
S10	26, 24	25
S11	26, 26	26
S12	26, 28	27
S13	28, 22	25
S14	28, 24	26
S15	28, 26	27
S16	28, 28	28

$\bar{x}$	Freq:	$P(\bar{x})$
22	1	0.0625
23	2	0.125
24	3	0.1875
25	4	0.25
26	3	0.1875
27	2	0.125
28	1	0.0625
total	16	1



An Estimate and Estimator

- Instead of asking the quality of an estimate, we will focus on the quality of the *estimation procedure*, or **estimator** ($\bar{X}$)
- Determine how good the estimator is by examining its **expected value, variance, and sampling distribution**

$\mu = 25; \sigma^2 = 5$

$\bar{x}$	Freq:	$P(\bar{x})$	$\bar{x}.P(\bar{x})$
22	1	0.0625	1.375
23	2	0.125	2.875
24	3	0.1875	4.5
25	4	0.25	6.25
26	3	0.1875	4.875
27	2	0.125	3.375
28	1	0.0625	1.75
		1	25

$E(\bar{x}) = \sum_{i=1}^{n} \bar{x}_i \, f(\bar{x}_i) = (22 \times 0.0625) + (23 \times 0.125) + (24 \times 0.1875) + (25 \times 0.25) + (26 \times 0.1875) + (27 \times 0.125) + (28 \times 0.0625) = 25 = \mu$

$$\begin{aligned}
\mathrm{var}(\bar{x}) &= E(\bar{x}^2) - \mu^2 \\
&= [(22^2 \times 0.0625) + (23^2 \times 0.125) + (24^2 \times 0.1875) + (25^2 \times 0.25) + (26^2 \times 0.1875) \\
&+ (27^2 \times 0.125) + (28^2 \times 0.0625)] - (25)^2 = 2.5
\end{aligned}$$

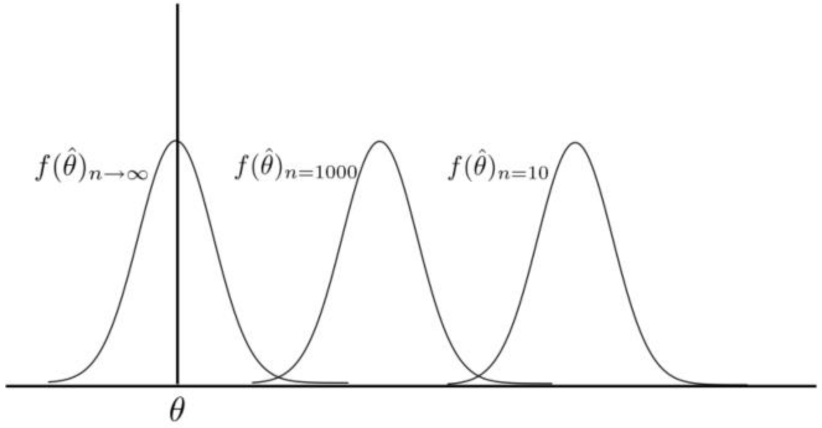
Properties of an Estimator (Unbiasedness)

- **Unbiasedness**: An estimator is said to be unbiased if the average of the estimates is equal to the true value of the population parameter

- Suppose $\hat{\theta}$ is an estimator for a parameter, θ

- $\hat{\theta}$ is an unbiased estimator if $E(\hat{\theta}) = \theta$

- If $E(\hat{\theta}) = \theta + b, \hat{\theta}$ is a biased estimator (i.e., $E(\hat{\theta}) \neq \theta$)
- if b can be reduced to zero when n becomes very large, $\hat{\theta}$ is described as **asymptotically unbiased**.



The Expected Value of Sample Means

- $\bar{x}$ can be written as $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n$
- The expected value of $\bar{X}$

$$E(\bar{x}) = E\left(\frac{1}{n}x_1\right) + E\left(\frac{1}{n}x_2\right) + \ldots + E\left(\frac{1}{n}x_n\right) \quad \because E(X+Y) = E(X) + E(Y)$$

$$= \frac{1}{n}E(x_1) + \frac{1}{n}E(x_2) + \ldots . + \frac{1}{n}E(x_n) \quad \because E(aX) = aE(X)$$

- Since x_i are also random variables $E(x_i)=\mu$, then

$$E(\bar{x}) = \frac{1}{n}\mu + \frac{1}{n}\mu + \ldots . + \frac{1}{n}\mu = \frac{1}{n}n\mu = \mu$$

Since $E(\bar{x}) = \mu$, the estimator $\bar{x}$ is unbiased.

The Expected Value of an estimator (Homework)

Suppose X is a random variable distributed as $X \sim N(\mu, \sigma^2)$

Check the following estimator $\tilde{\theta}$ for population mean μ is an unbiased estimator.

$$\tilde{\theta} = \frac{\sum_{i=1}^{n} x_i}{n+3}$$

Is the estimator $\tilde{\theta}$ an unbiased estimator?

The Variance of Sample Means

With random sampling the observations, x_i are statistically independent, and thus are uncorrelated

$$\bar{X} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n$$

$$\text{var}(\bar{X}) = \text{var}\left(\frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n\right)$$

$$= \frac{1}{n^2}\text{var}(x_1) + \frac{1}{n^2}\text{var}(x_2) + \ldots. + \frac{1}{n^2}\text{var}(x_n)$$

$$\because \text{var}(aX) = a^2 \text{ var}(X)$$

Since x_i are identically distributed, $\text{var}(x_i) = \sigma^2$

$$= \frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \ldots. + \frac{1}{n^2}\sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Hence, the variance of sample mean is smaller than the population variance if n>1

The Variance of an estimator (Homework)

Suppose X is a random variable distributed as $X \sim N(\mu, \sigma^2)$

Find the variance of the following estimator $\tilde{\theta}$

$$\tilde{\theta} = \frac{\sum_{i=1}^{n} x_i}{n+3}$$

Estimating the Variance of Sample Means

$$\text{var}(\bar{X}) = \frac{\sigma^2}{n}$$

$$\hat{\sigma}^2 = \frac{\sum(x_i - \bar{X})^2}{n-1}$$

- Since the population variance, σ^2, is an unknown parameter, we can use sample variance, $\hat{\sigma}^2$, as an estimator of the population variance
- Using the sample variance, we can estimate the variance of the estimator sample mean, $\bar{X}$ as:

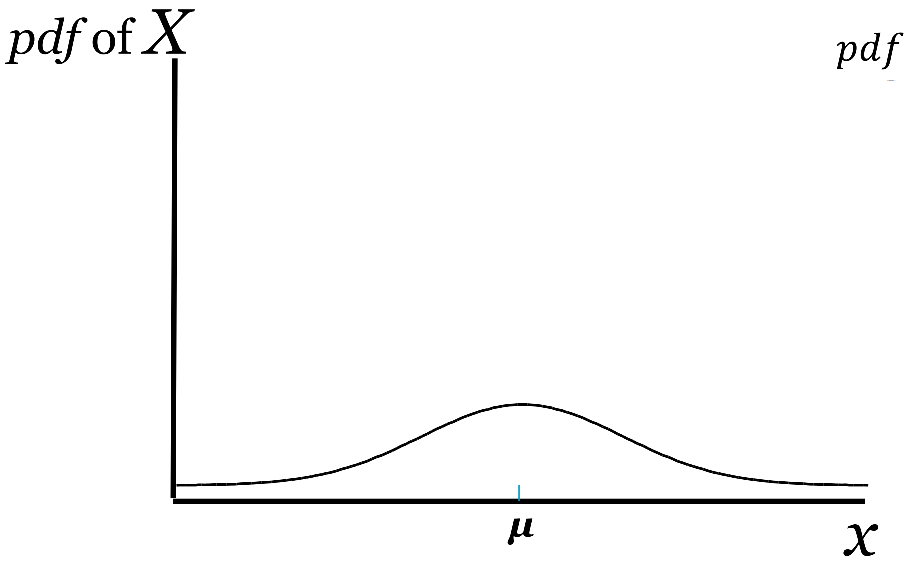
$$\widehat{\text{var}}\left(\bar{X}\right) = \frac{\hat{\sigma}^2}{n}$$

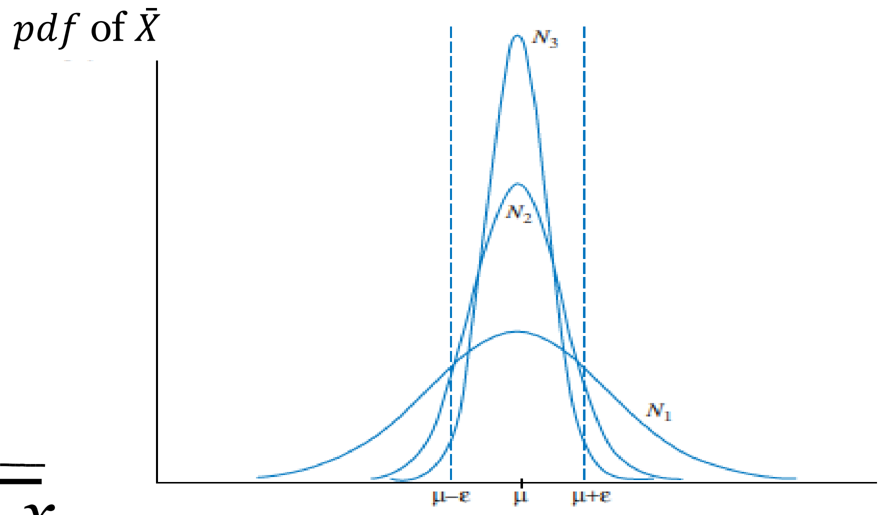
- The square root of the estimated variance is called the **standard error** of $\bar{X}$ and is also known as the **standard error of the estimate**:
- $\widehat{\text{s.e}}(\bar{X}) = \sqrt{\widehat{\text{var}}\ (\bar{X})} = \frac{\hat{\sigma}}{\sqrt{n}}$

The larger the sample size, the smaller the variance of sample mean, hence the standard error of the estimate.

Distribution of Sample Means

Increasing sample size and sampling distributions of sample means ($\bar{X}$)





The distribution of sample mean is more compact. Why?

The larger the sample size, the smaller the variance of sample mean.

What will happen if sample size n converges to the size of population, i.e., $n \to \infty$?

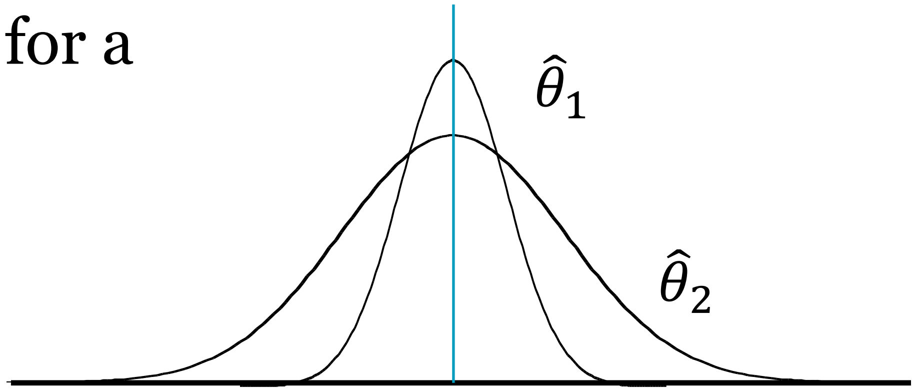
Properties of an Estimator (Efficiency)

Efficiency of an estimator

In a class of unbiased estimators calculated using the same sample size, the one with the smallest variance is called the **minimum variance unbiased estimator** or **best estimator** or **efficient estimator**.

Suppose, we have two unbiased estimator, $\hat{\theta}_1$ and $\hat{\theta}_2$ for a population parameter θ

$$E(\hat{\theta}_1) = \theta$$
$$E(\hat{\theta}_2) = \theta$$
$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$$



$\hat{\theta}_1$ which has a smaller variance is an ~~(more)~~ efficient estimator

Efficiency of an Estimator (Quiz)

We employ two estimators, Z and sample mean to predict the population mean, μ of random variable X.

Generalized estimator $Z = \lambda_1 X_1 + \lambda_2 X_2$, where λ_1 =0.8 and λ_2 =0.2. Is this estimator unbiased?

$$Var(Z) = (\lambda_1^2 + \lambda_2^2)\sigma^2$$

Sample mean $\bar{X} = \dfrac{\sum x_i}{n}$ where n = 2.

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

Which one is an ~~(more)~~ efficient estimator?

Properties of an Estimator (Consistency)

Consistency of an estimator

An estimator which is unbiased (or asymptotically unbiased) is said to be consistent if its variance reduces to zero when the sample size becomes very large.

$$E(\hat{\theta}) = \theta \text{ or } E(\hat{\theta}) = \theta \text{ when } n \to \infty$$

$$\text{var}(\hat{\theta}) = 0 \text{ when } n \to \infty$$

Is the estimator sample mean a consistent estimator?

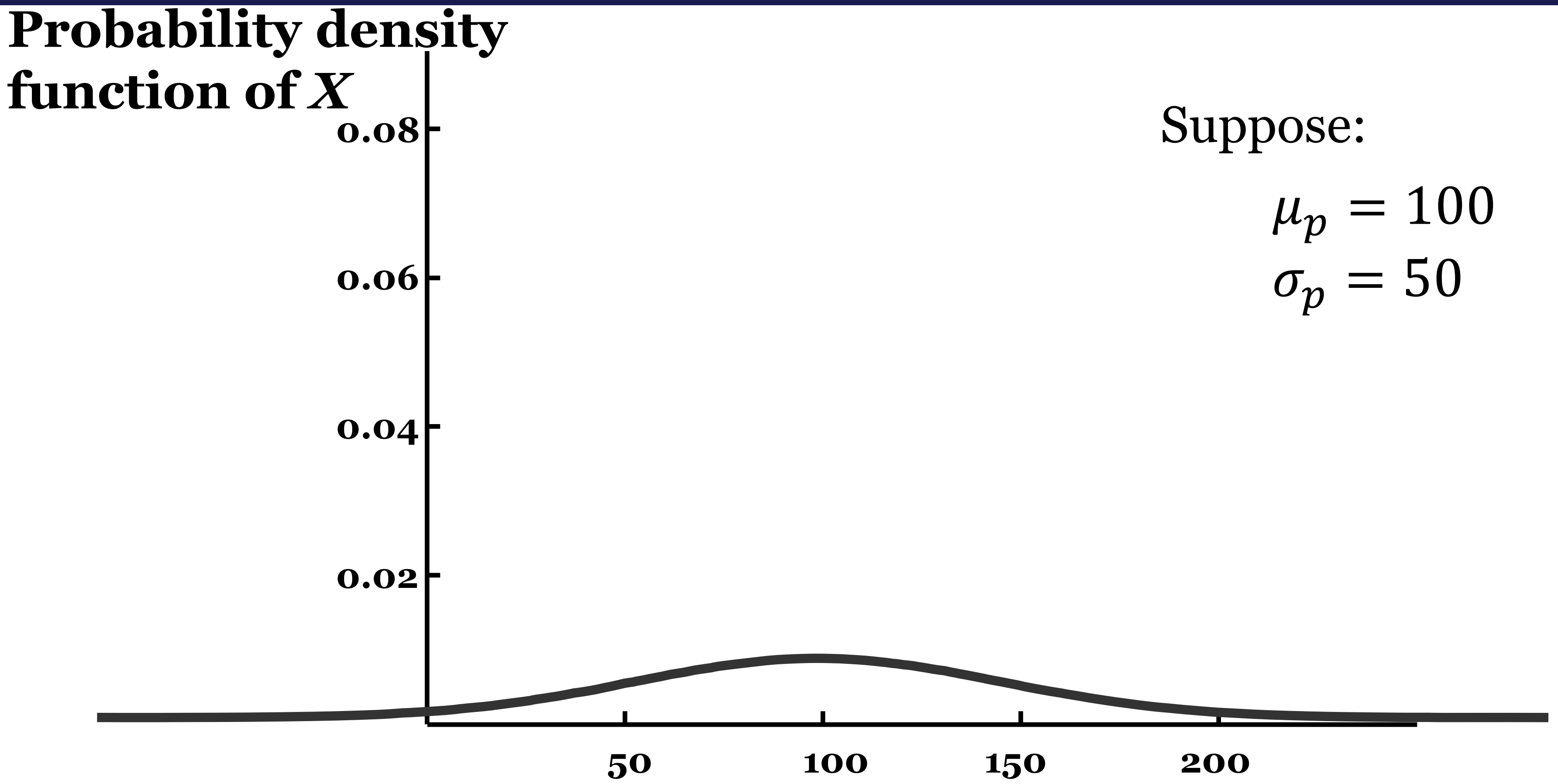
The consistency of an estimator (Homework)

Suppose X is a random variable distributed as $X \sim N(\mu, \sigma^2)$

Check whether the following estimator $\tilde{\theta}$ is a consistent estimator.

$$\tilde{\theta} = \frac{\sum_{i=1}^{n} x_i}{n+3}$$

Increasing sample size & variance of sample mean

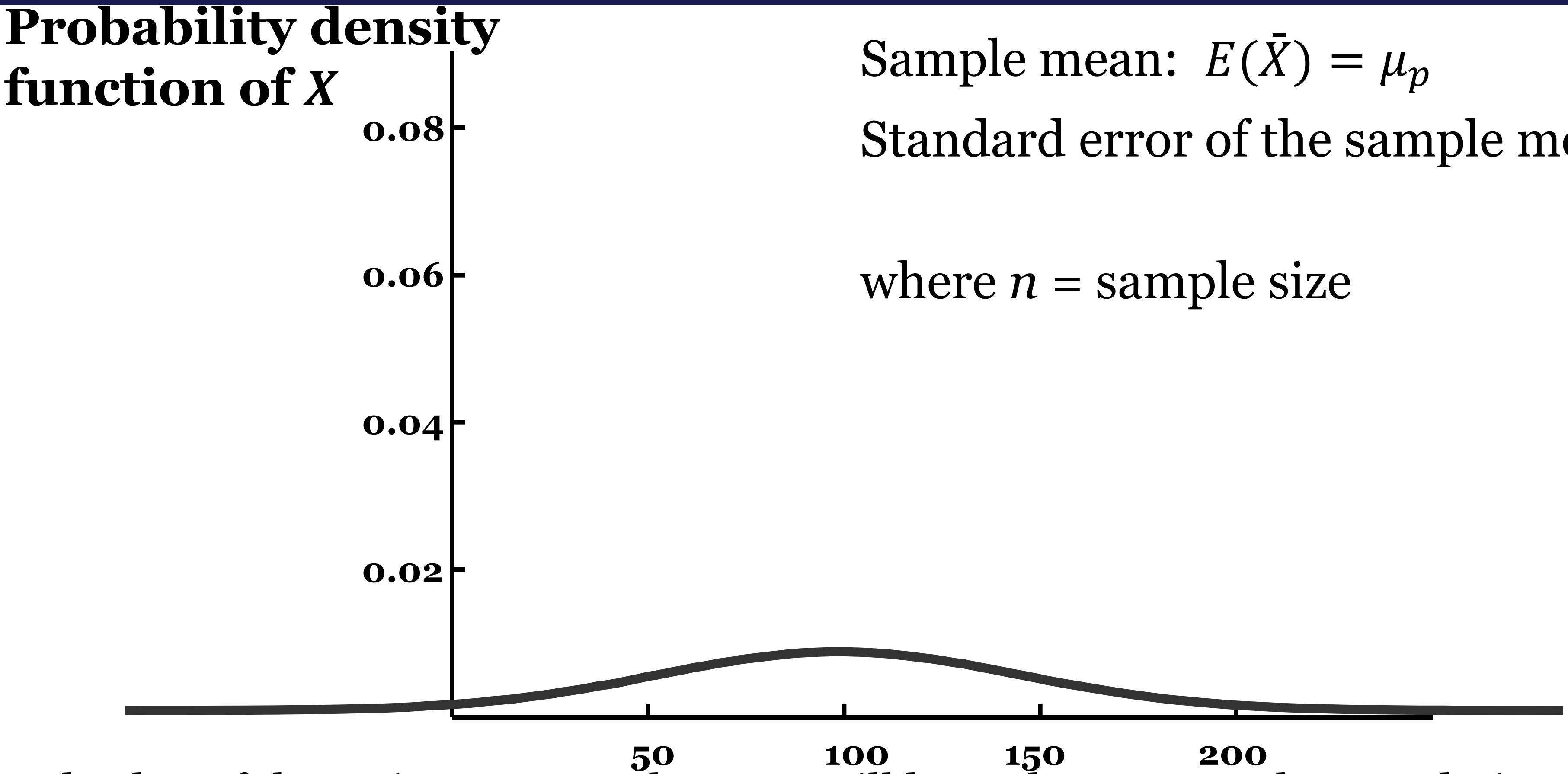


Suppose:

$$\mu_p = 100$$
$$\sigma_p = 50$$

Suppose that a random variable *X* has population mean 100 and standard deviation 50, as in the diagram. Suppose that we <u>do not know</u> the population mean and we are using the sample mean to estimate it.

Increasing sample size & variance of sample mean



Sample mean: $E(\bar{X}) = \mu_p$

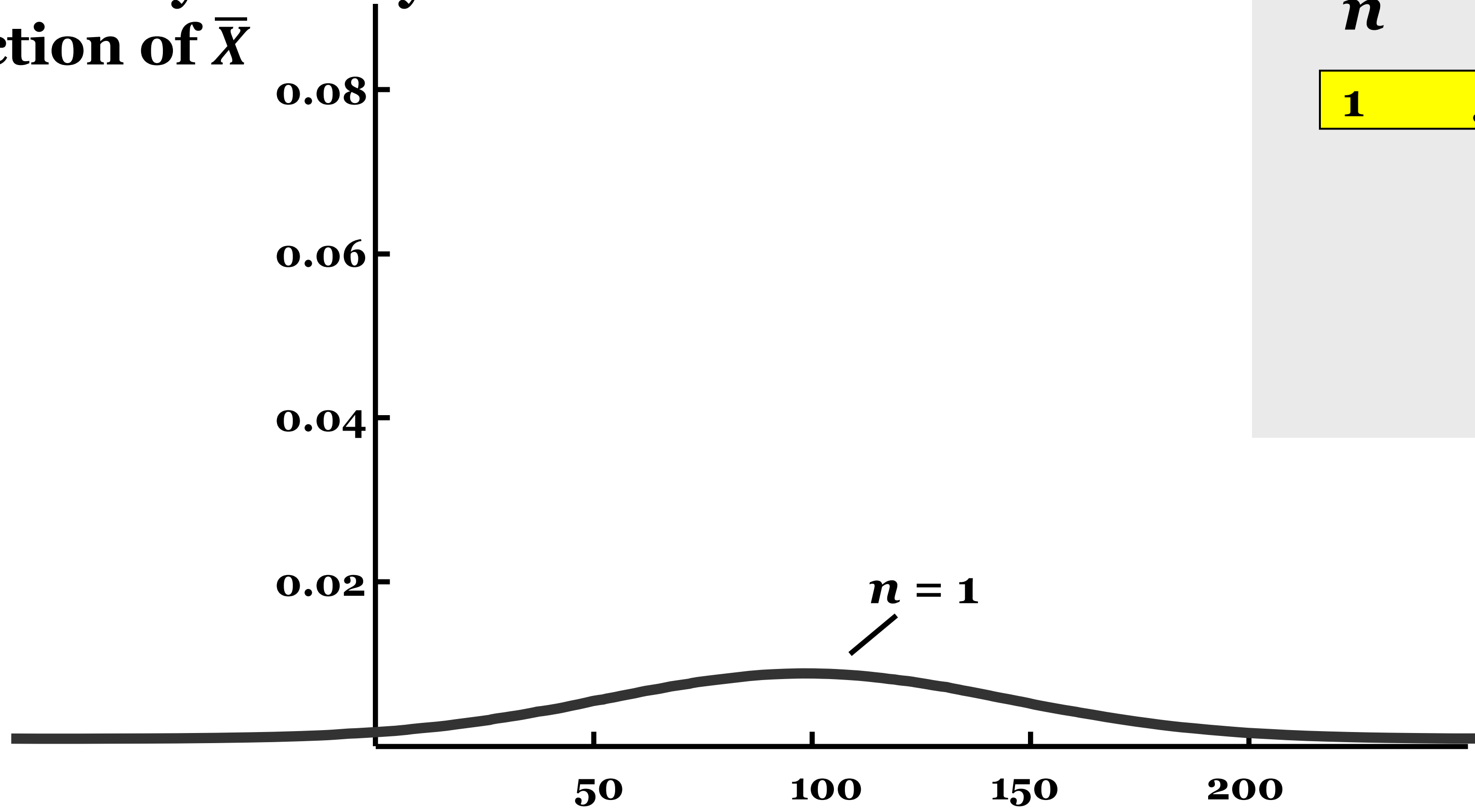
Standard error of the sample mean: $\frac{\hat{\sigma}}{\sqrt{n}}$

where n = sample size

The expected value of the estimator sample mean will have the same as the population mean X, but its standard deviation will be $50/\sqrt{n}$, where n is the number of observations in the sample. The larger is the sample, the smaller will be the standard deviation of the sample mean.

Increasing sample size & variance of sample mean

Probability density function of $\bar{X}$

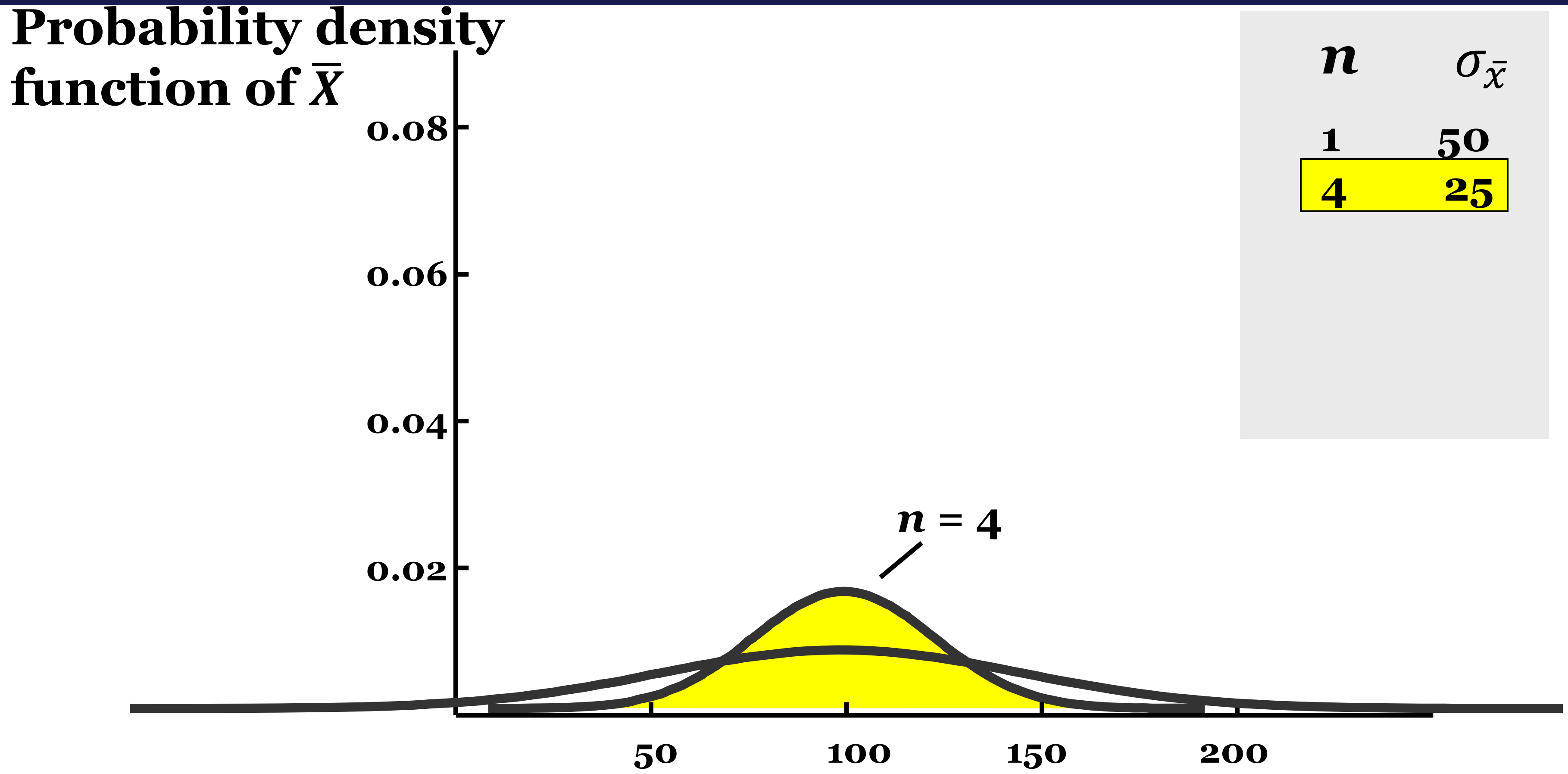


n	$\sigma_{\bar{x}}$
1	50

$$\mu_p = 100$$
$$\sigma_p = 50$$
$$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$$

If n is equal to 1, the sample consists of a single observation. The distribution of $\bar{X}$ is the same as X and its standard deviation is 50.

Increasing sample size & variance of sample mean

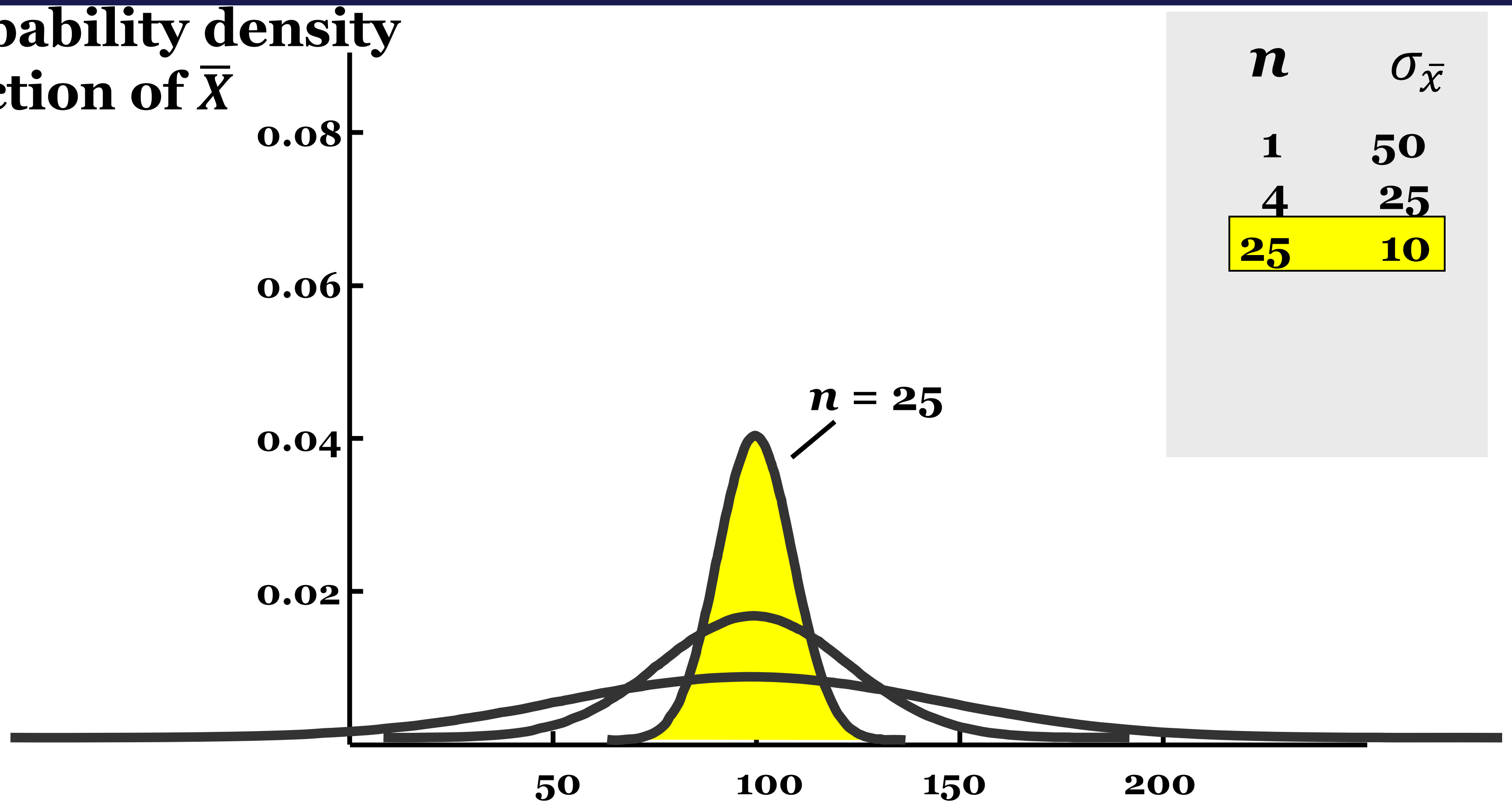


n	$\sigma_{\bar{x}}$
1	50
4	25

$$\mu_p = 100$$
$$\sigma_p = 50$$
$$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$$

We will see how the shape of the distribution changes as the sample size is increased.

Increasing sample size & variance of sample mean

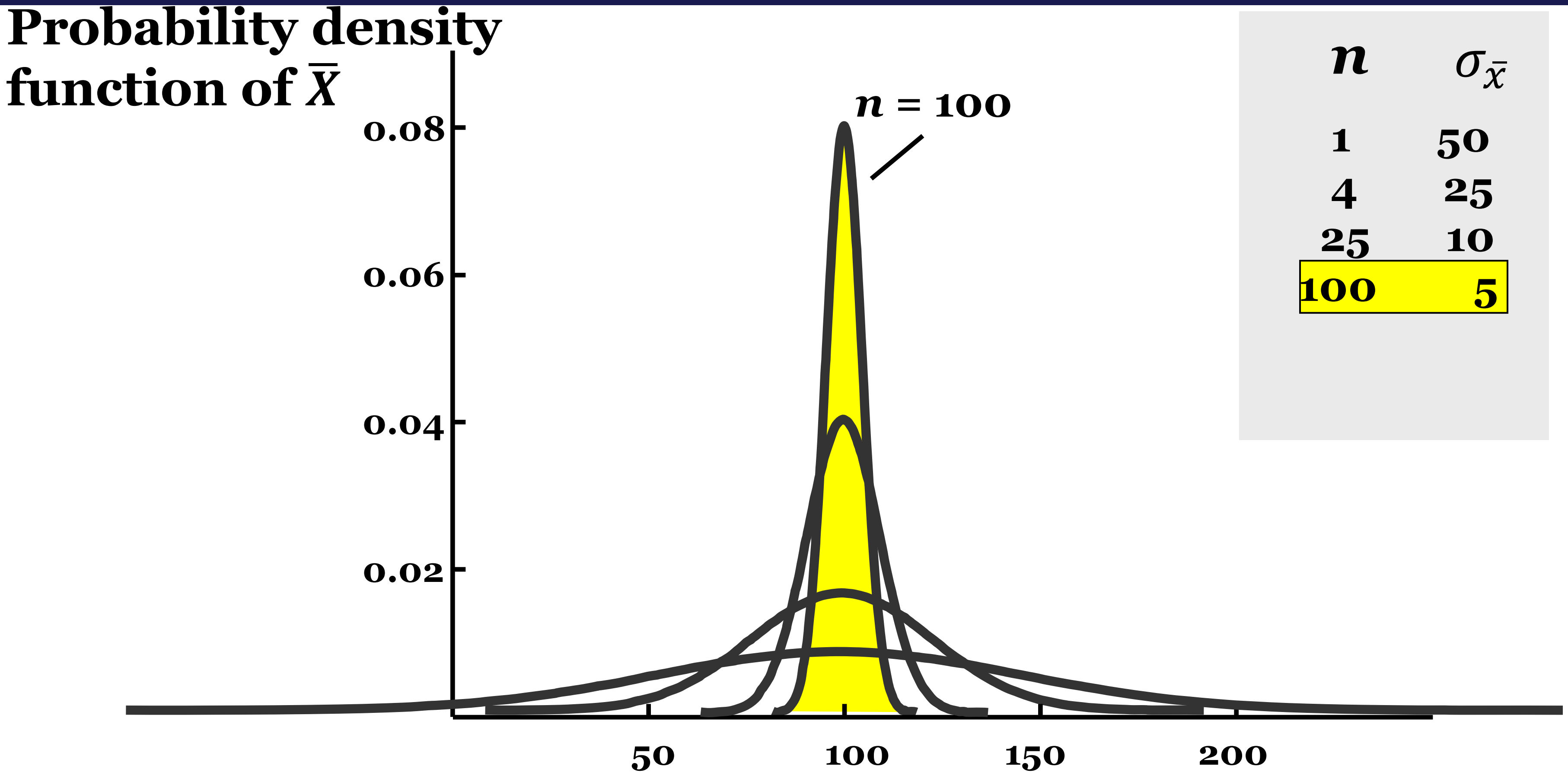


n	$\sigma_{\bar{x}}$
1	50
4	25
25	10

$$\mu_p = 100$$
$$\sigma_p = 50$$
$$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$$

The distribution becomes more concentrated about the population mean

Increasing sample size & variance of sample mean

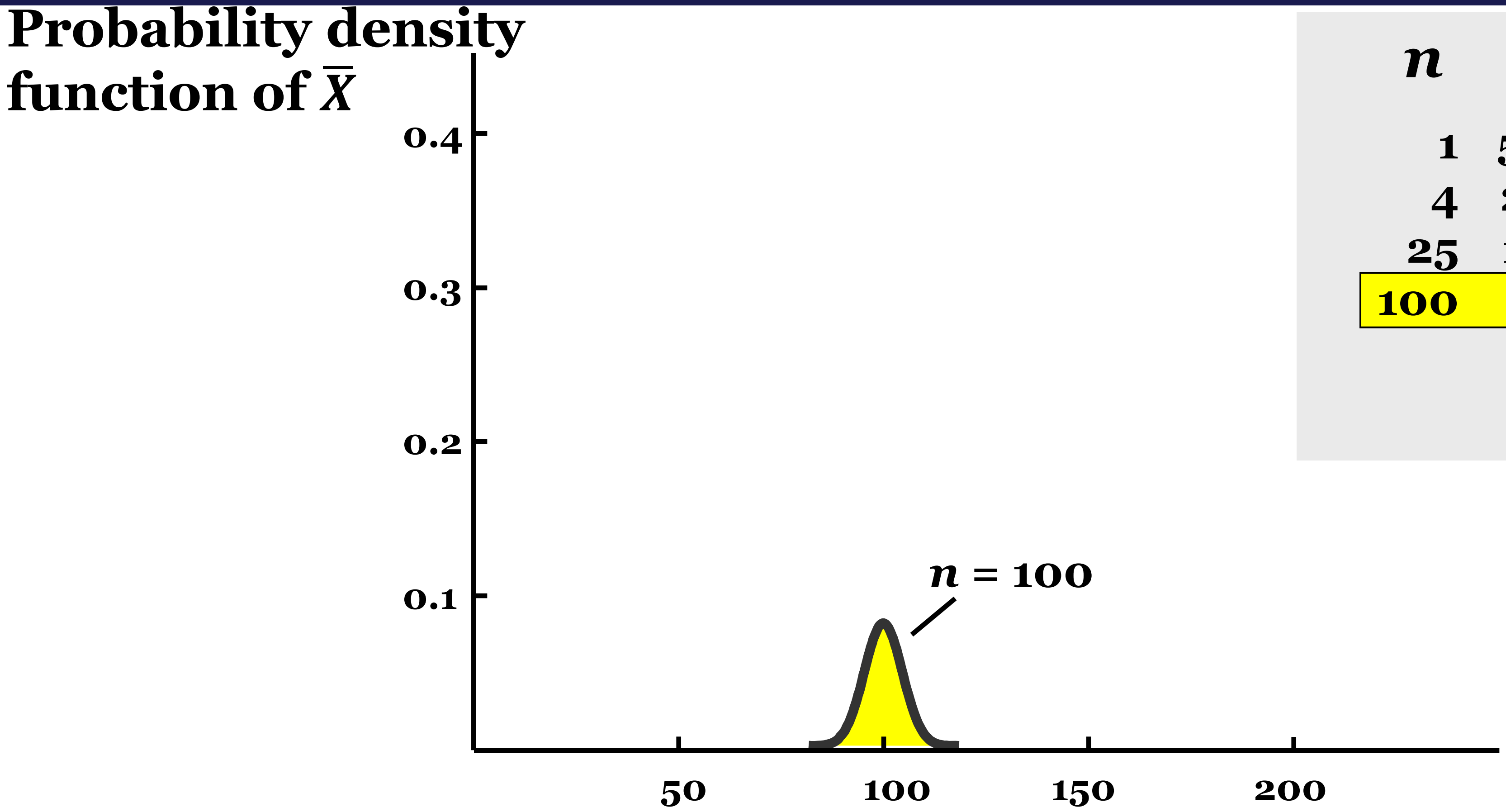


n	$\sigma_{\bar{x}}$
1	50
4	25
25	10
100	5

$\mu_p = 100$
$\sigma_p = 50$
$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$

To see what happens when n is greater than 100, we will have to change the vertical scale from 50:1 to 10:1.

Increasing sample size & variance of sample mean



n	$\sigma_{\bar{x}}$
1	50
4	25
25	10
100	5

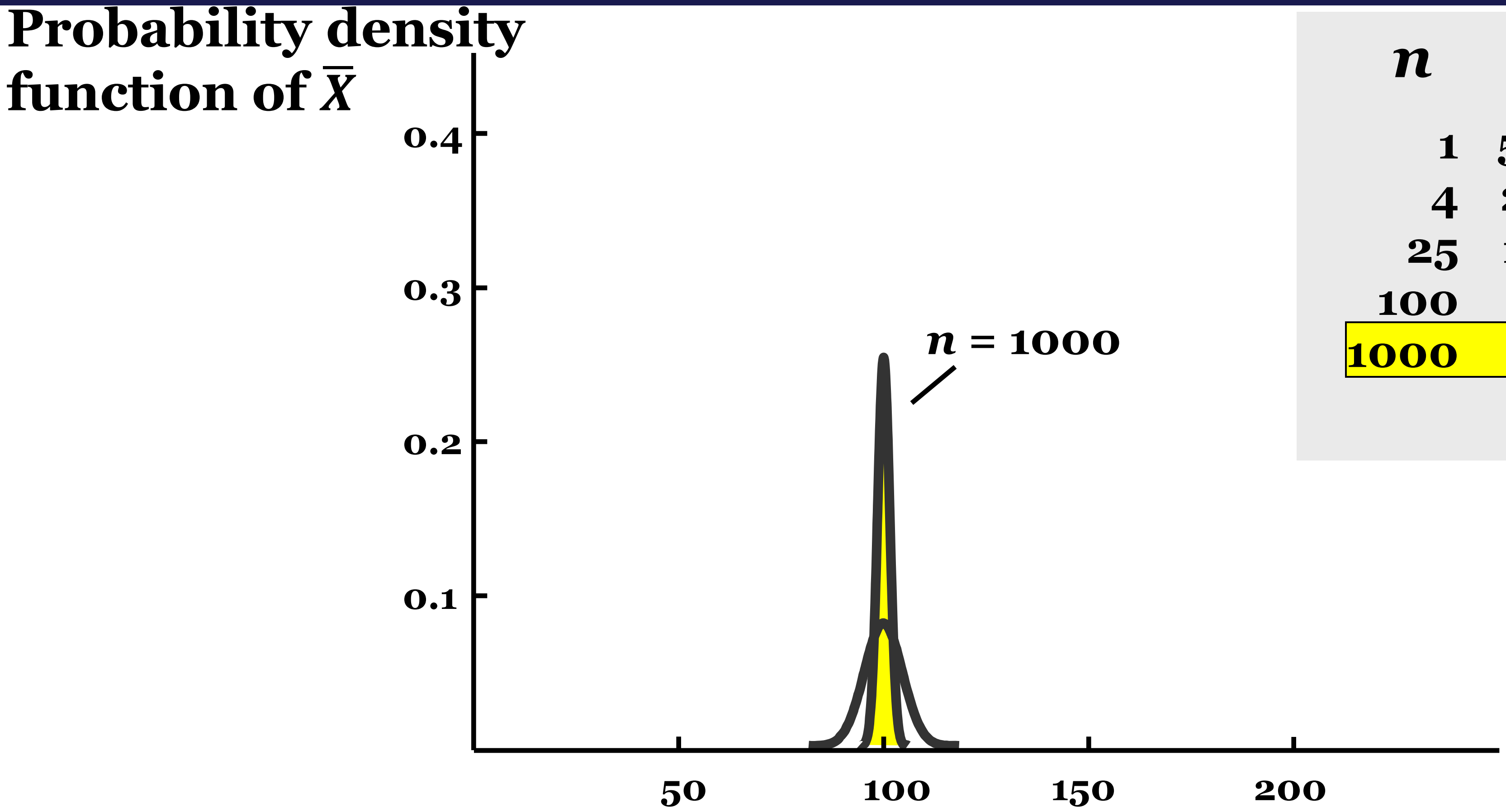
$\mu_p = 100$

$\sigma_p = 50$

$\sigma_{\bar{x}} = \dfrac{\sigma_p}{\sqrt{n}}$

We have reduced the vertical scale by a factor of 5.

Increasing sample size & variance of sample mean



n	$\sigma_{\bar{x}}$
1	50
4	25
25	10
100	5
1000	1.6

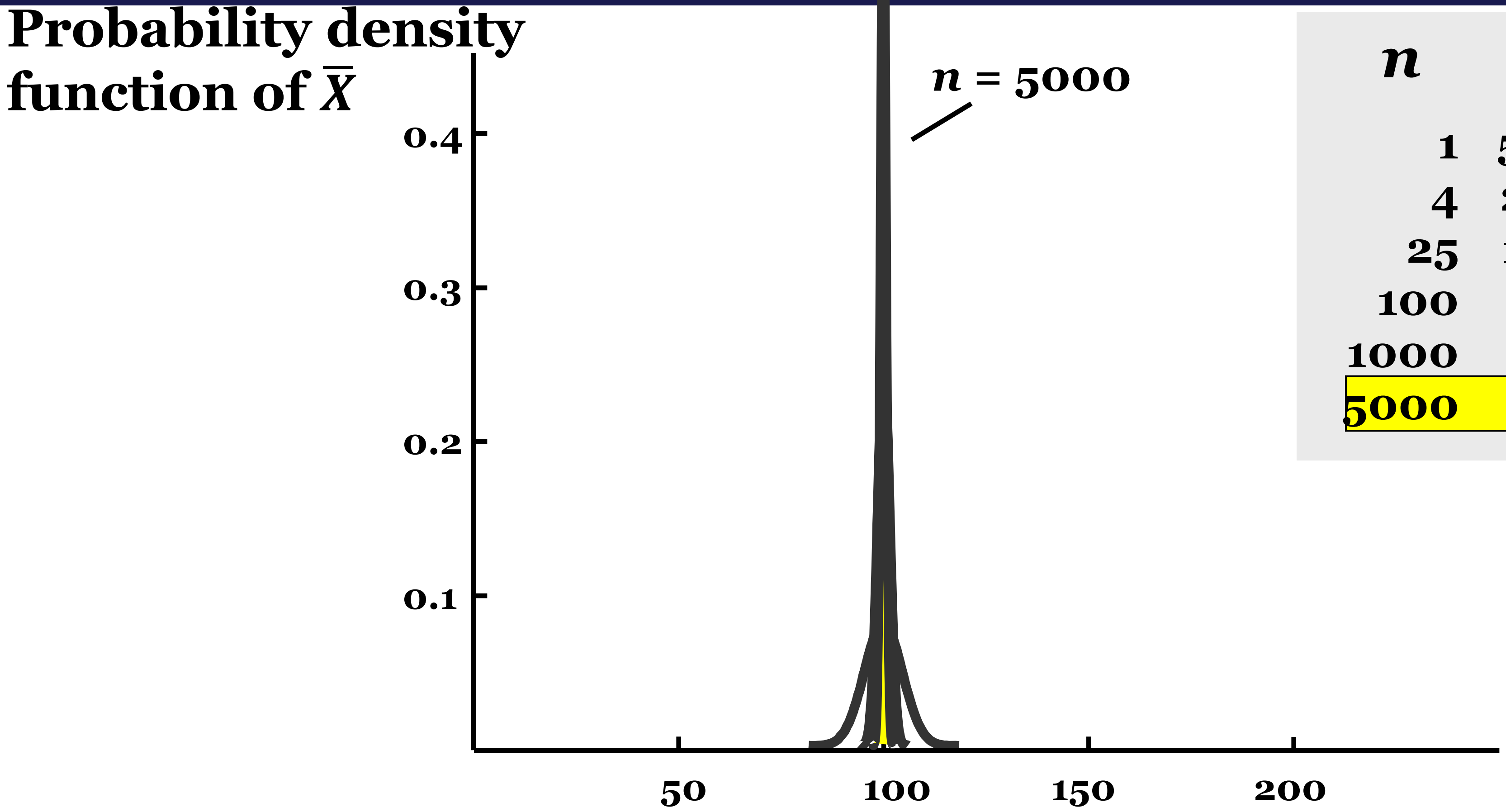
$\mu_p = 100$

$\sigma_p = 50$

$$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$$

The distribution continues to contract about the population mean.

Increasing sample size & variance of sample mean



n	$\sigma_{\bar{x}}$
1	50
4	25
25	10
100	5
1000	1.6
5000	0.7

$\mu_p = 100$

$\sigma_p = 50$

$$\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$$

When n approaches infinity, the variance of the distribution tends to zero. The distribution collapses to a spike at the true value. The sample mean is therefore a consistent estimator of the population mean.

Effect of increasing the sample size

Finite samples: $\bar{X}$ is an unbiased estimator of μ ($\mathrm{E}(\bar{X}) = \mu$)

Large samples: the probability distribution of $\bar{X}$ collapses to a spike at μ

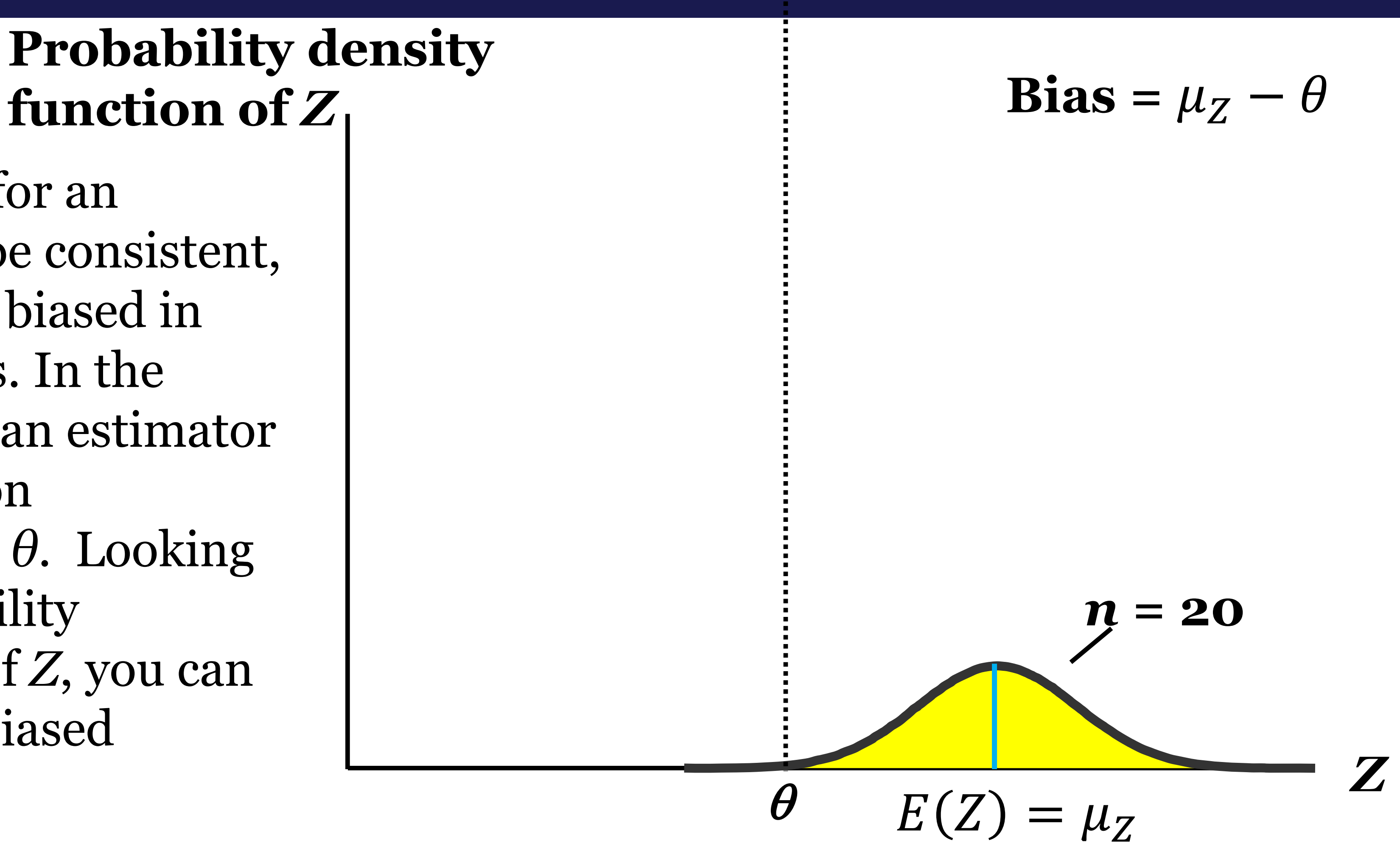
plim $\bar{X} = \mu$ (Probability Limit)

Consistency is a large-sample concept. A consistent estimator becomes an increasingly accurate estimator of the population characteristic, and in the limit, the consistent estimator equal to it.

As the sample size becomes extremely large, the distribution of the sample mean collapses to a spike located at the true value. The sample mean is therefore consistent as well as unbiased.

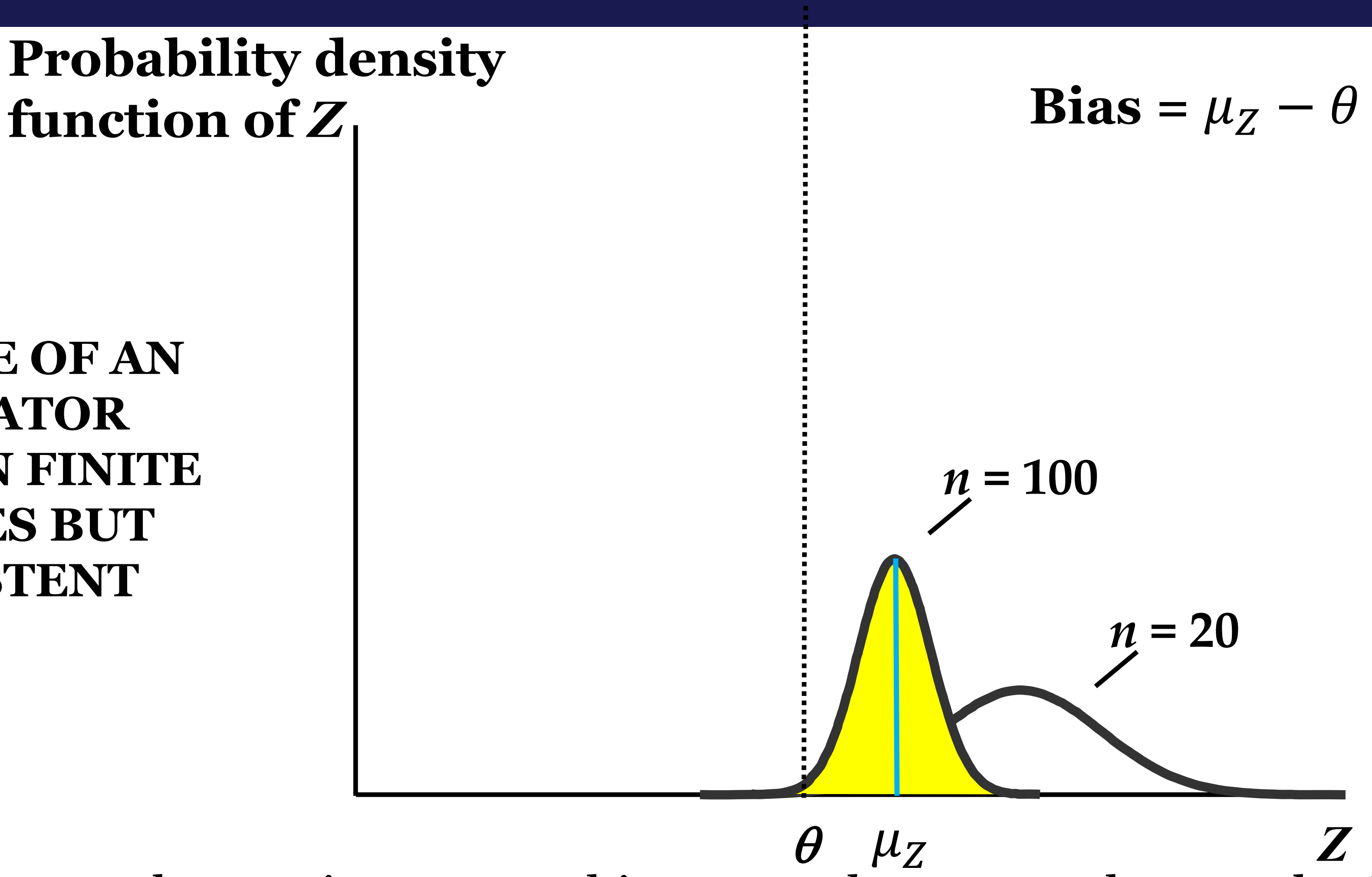
Increasing sample size & asym. unbiasedness

It is possible for an estimator to be consistent, despite being biased in finite samples. In the diagram, Z is an estimator of a population characteristic θ. Looking at the probability distribution of Z, you can see that Z is biased upwards.



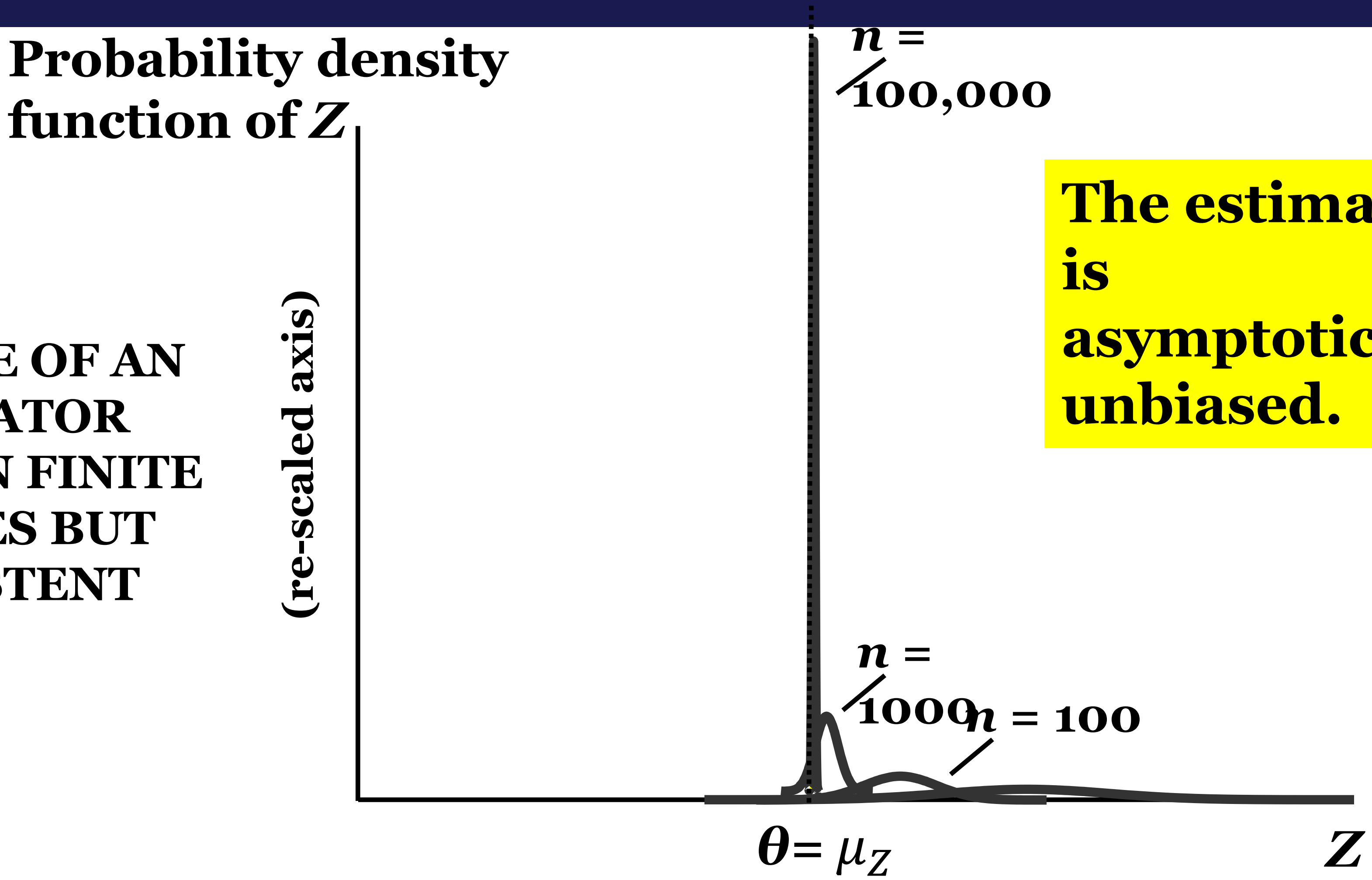
Increasing sample size & asym. unbiasedness

EXAMPLE OF AN ESTIMATOR BIASED IN FINITE SAMPLES BUT CONSISTENT



For the estimator to be consistent, <u>two things must happen</u> as the sample size increases. One is that the bias should diminish as n increases, as shown here.

Increasing sample size & asym. unbiasedness



EXAMPLE OF AN ESTIMATOR BIASED IN FINITE SAMPLES BUT CONSISTENT

The estimator is asymptotically unbiased.

The other is that the distribution should collapse to a spike.
In this case, both conditions are approximately satisfied when n=100,000.

Summary

After this lecture, you should be able to:

- Understand the concept of Statistical Inference
- Differentiate between an estimate and an estimator
- Calculate the expected value and variance of estimates
- Identify the Properties of an Estimator
    - Unbiasedness
    - Efficiency
    - Consistency

Thank you and any question?