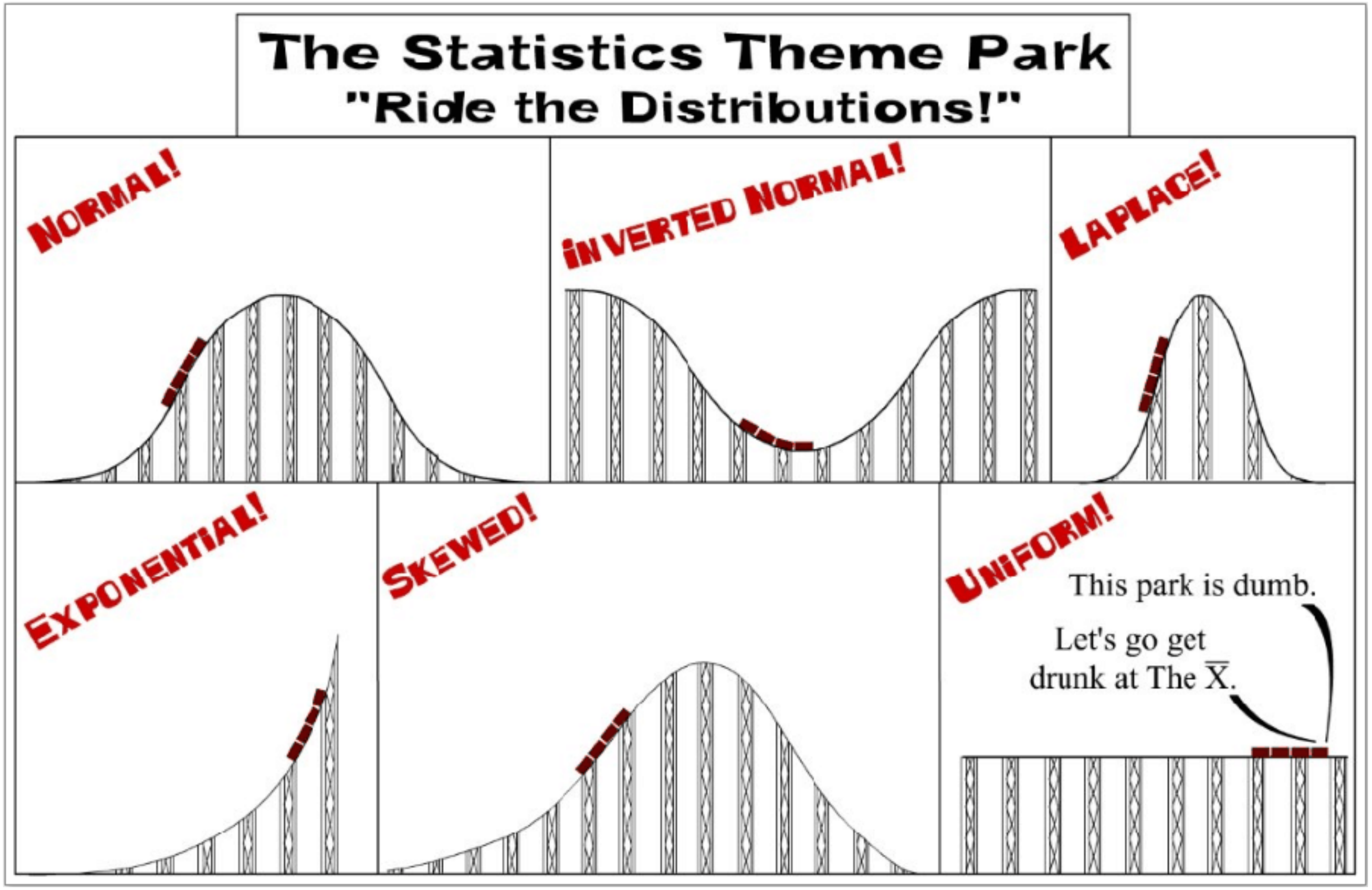
CS 237: Probability in Computing

Instructors: Alina Ene, Tiago Januario



BOSTON UNIVERSITY

CONTINUOUS DISTRIBUTIONS

- Several of the common discrete distributions have continuous analogues
- Binomial vs. Normal distribution
  - Number of successes in a finite number of IID trials → discrete, Binomial
    - Produces a symmetric “bell-shaped” PDF
  - Analogue with continuous data: Normal distribution

BOSTON UNIVERSITY

CONTINUOUS DISTRIBUTIONS

- Several of the common discrete distributions have continuous analogues
- Geometric vs. Exponential distribution
  - Number of IID trials until an event occurs → discrete, Geometric
  - Amount of **time** until an event occurs → continuous, Exponential

CONTINUOUS DISTRIBUTIONS

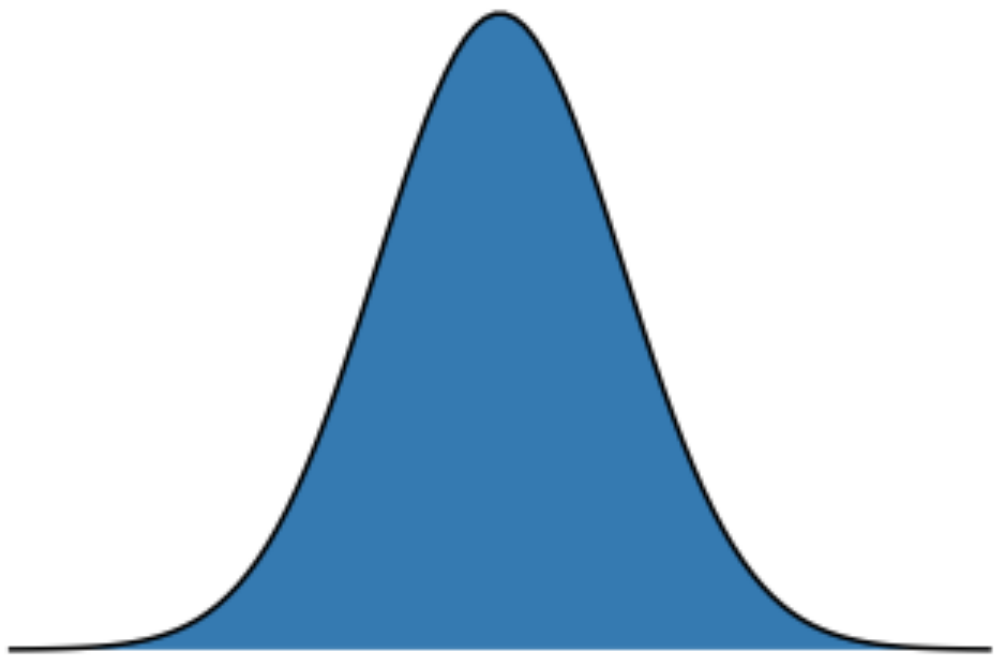
- Several of the common discrete distributions have continuous analogues
- Binomial vs. Normal distribution
- Geometric vs. Exponential distribution
- Normal and Exponential can be seen as **limits** of their discrete analogues (more on this shortly)

BOSTON UNIVERSITY

NORMAL DISTRIBUTION

- Very familiar to us as the "bell curve"
- Models a wide range of quantities of interest:
  - Measurement error
  - Height of individuals
  - Averages of many independent values (central limit thm)
  - Test scores (e.g., SAT, IQ)
  - Pizza delivery times
  - …

NORMAL DISTRIBUTION



- Range: $\mathbb{R}$
- Parameters: μ, σ^2 (expectation and variance)
- Notation: $\text{Normal}(\mu, \sigma^2)$
- PDF and CDF:

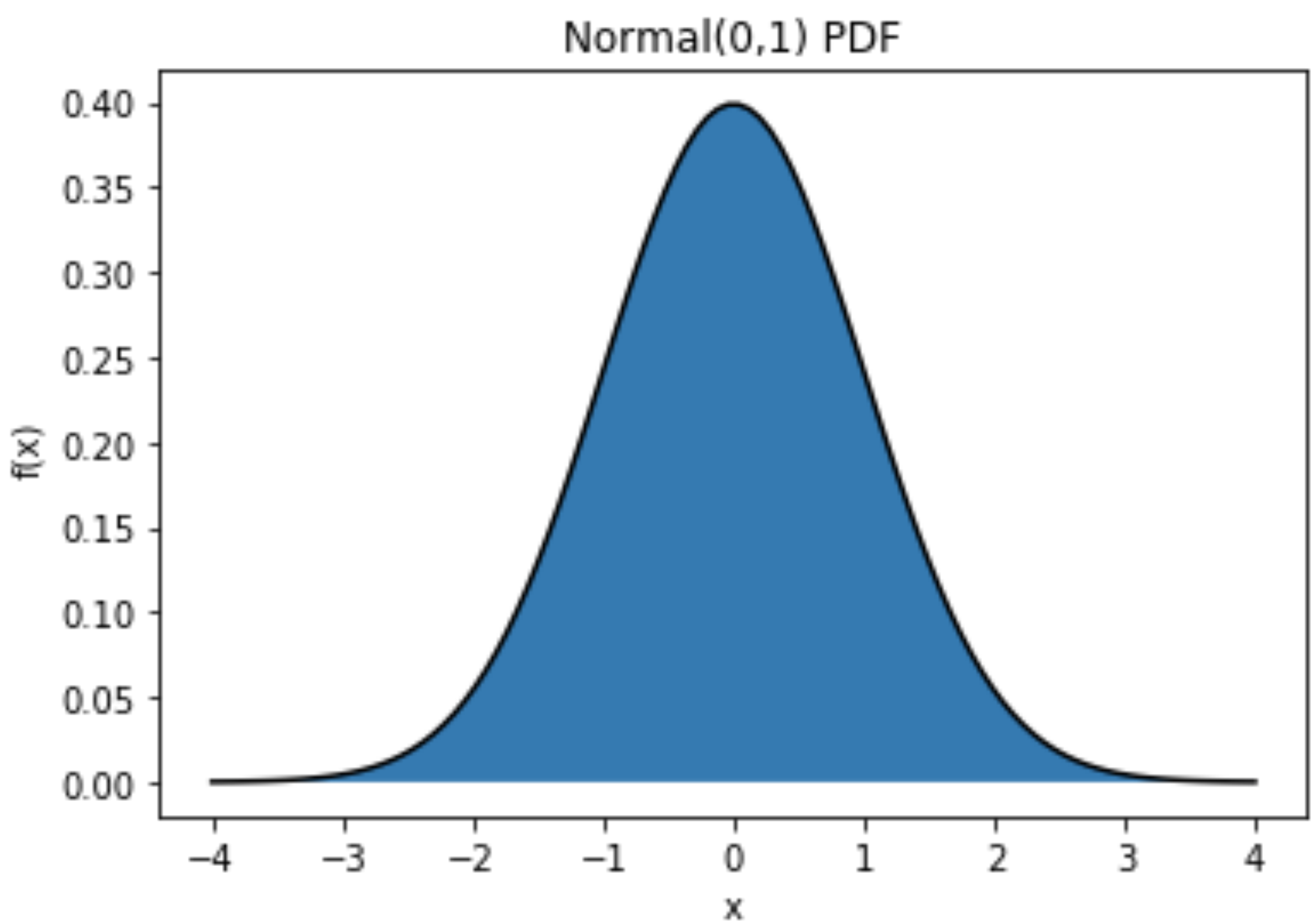
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

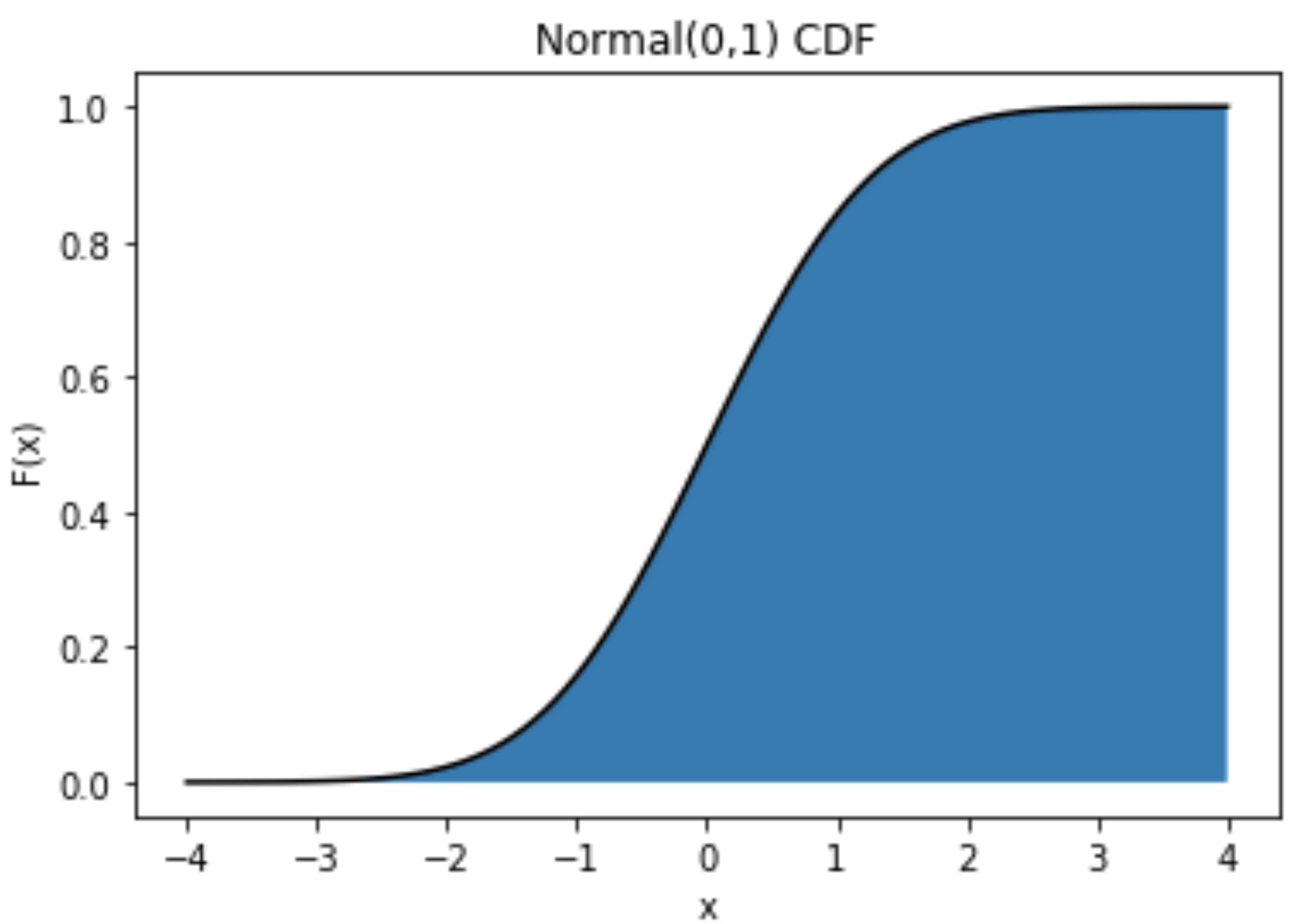
$$F(x) = \int_{-\infty}^{x} f(t)dt$$ (no closed-form formula exists)

BOSTON UNIVERSITY

NORMAL DISTRIBUTION

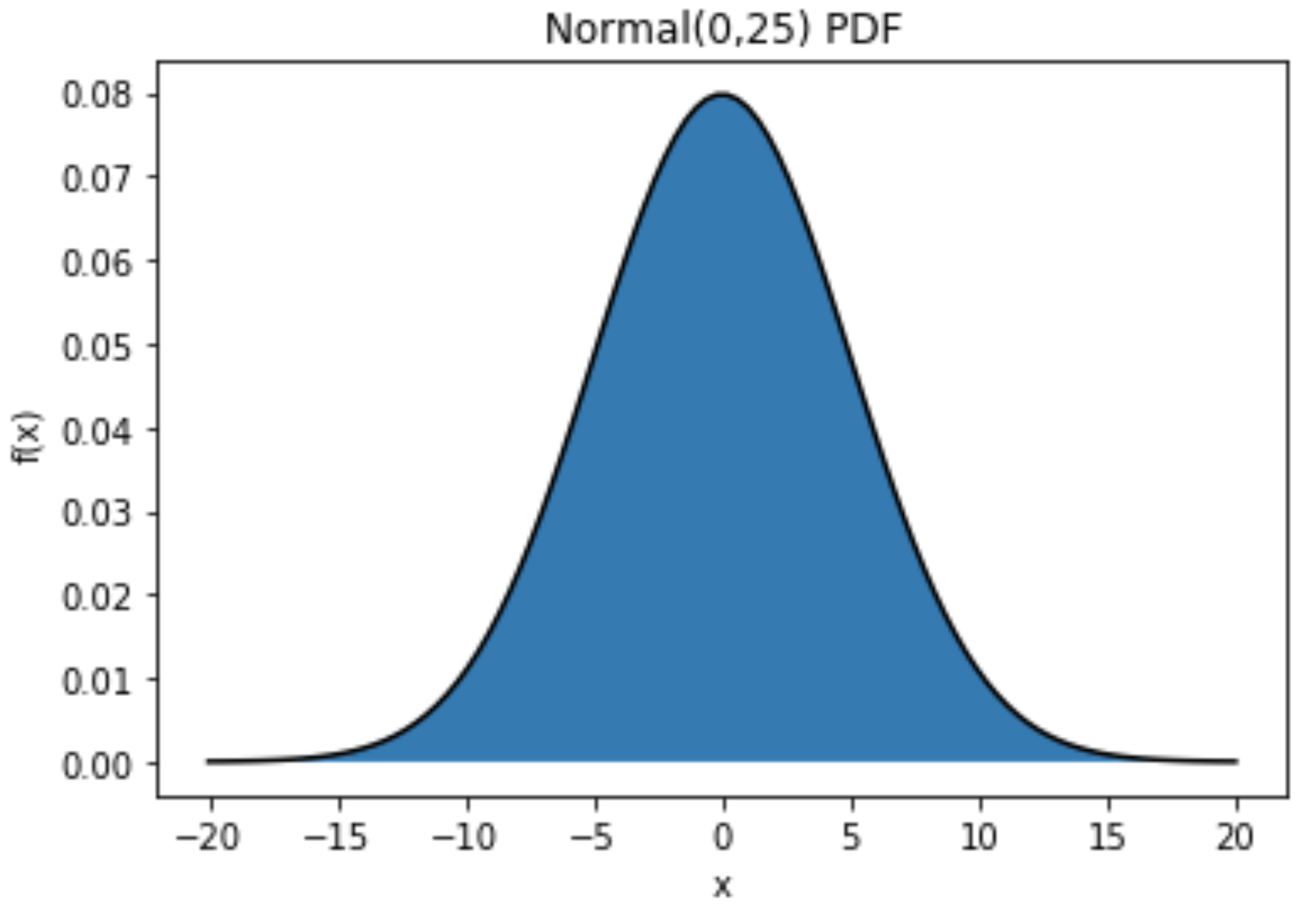
$$\text{Normal}(\mu, \sigma^2) \quad \text{PDF:} \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

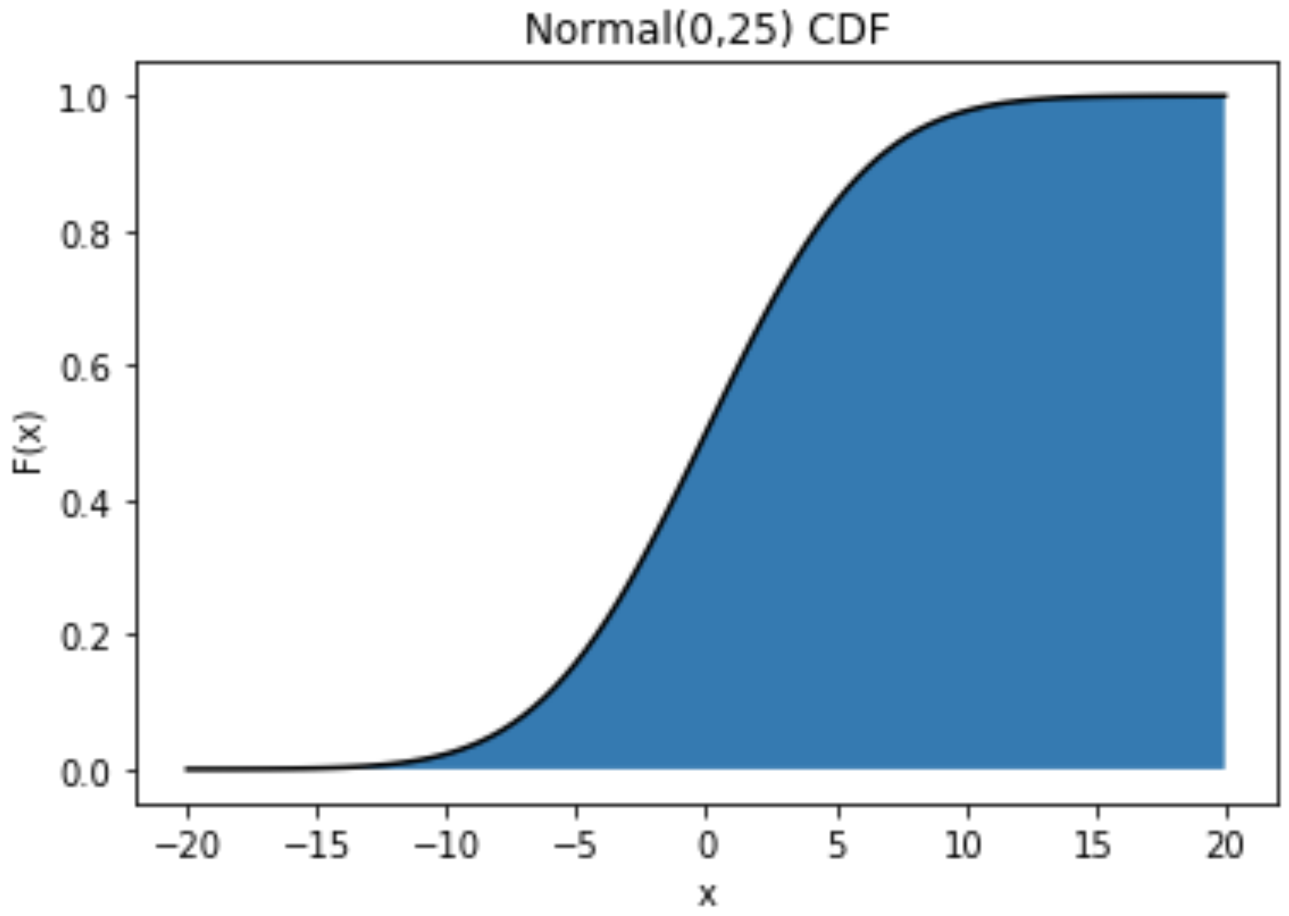




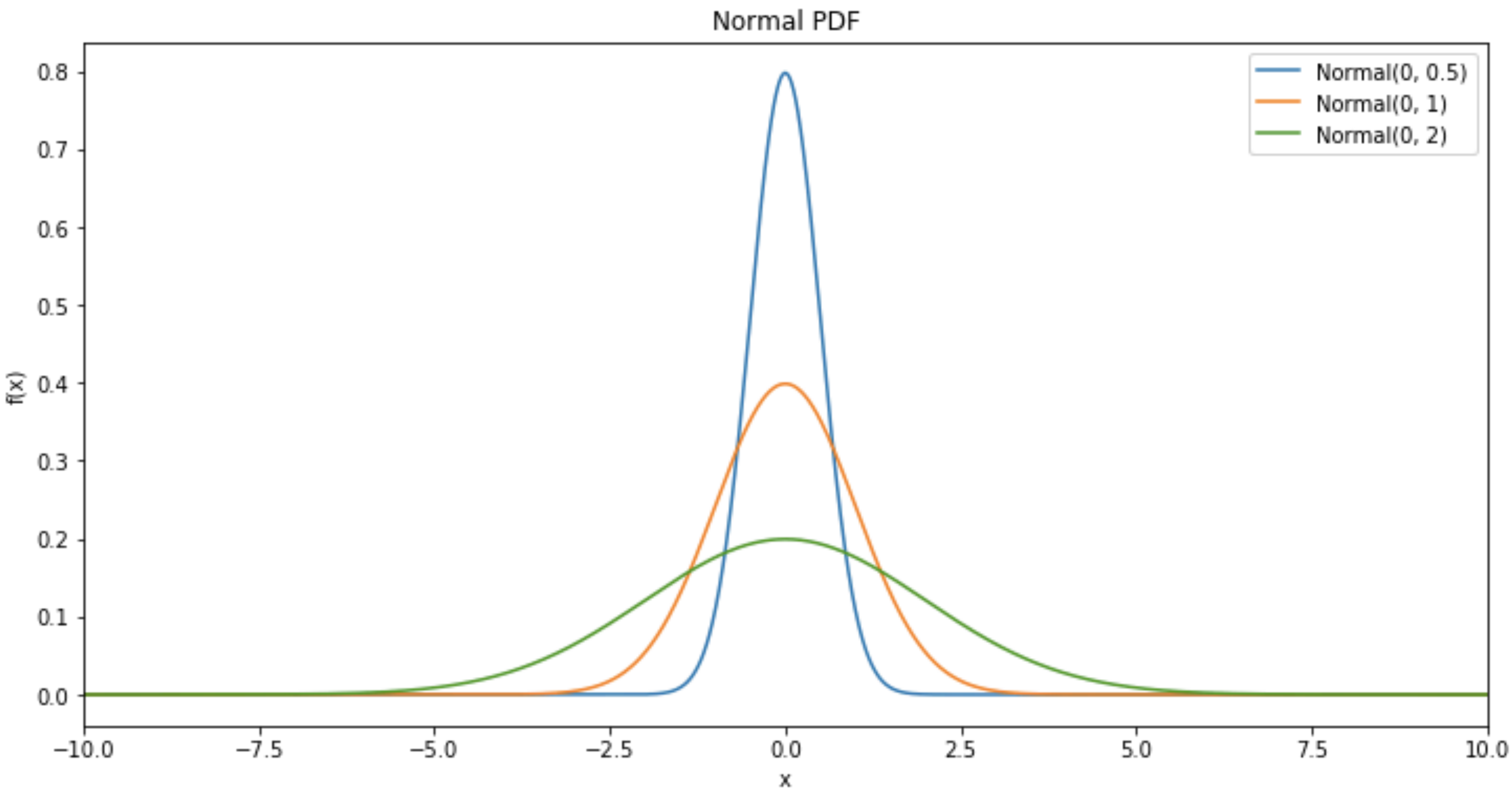
NORMAL DISTRIBUTION

$$\text{Normal}(\mu, \sigma^2) \quad \text{PDF:} \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

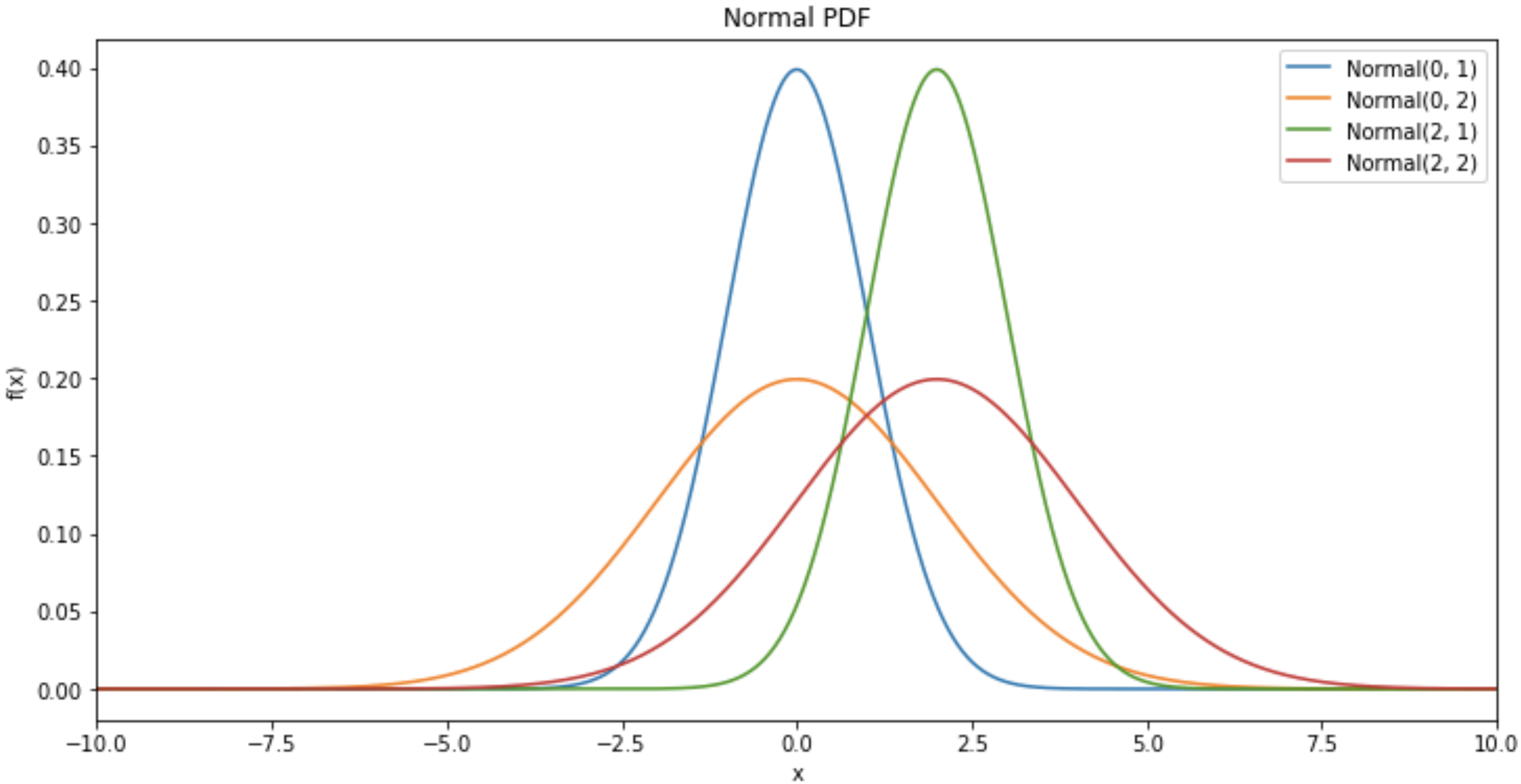




NORMAL DISTRIBUTION



NORMAL DISTRIBUTION



STANDARD NORMAL DISTRIBUTION

- Normal(0, 1) is called the standard normal distribution
- It arises so often that its PDF and CDF have special notation

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$ Normal(0, 1) PDF

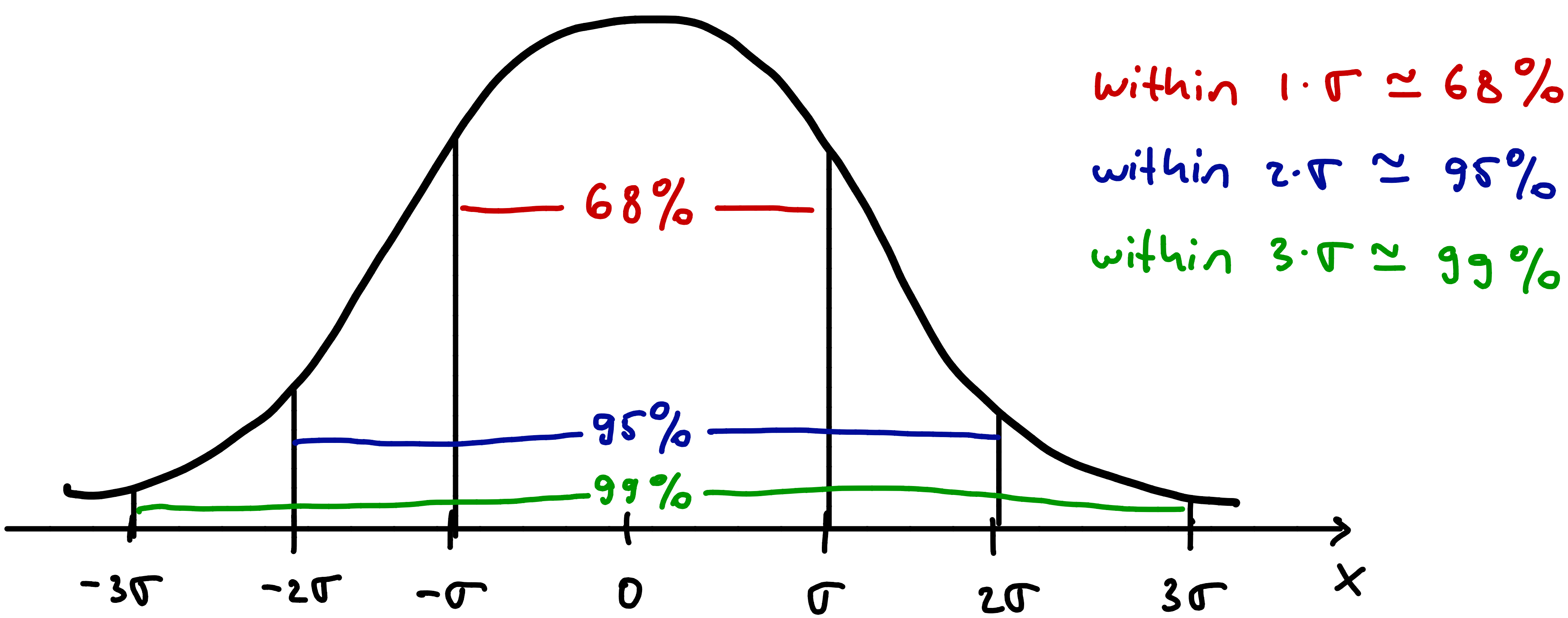
$$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt$$ Normal(0, 1) CDF

- It is also common to use the letter Z to denote a Normal(0,1) random variable

BOSTON UNIVERSITY

STANDARD NORMAL DISTRIBUTION

- It is useful to remember the following probabilities:



STANDARD NORMAL CDF

- Since the Normal CDF does not have a closed-form formula, we need some way to evaluate it
- There are several options, as we illustrate next:
  - Look up the standard Normal CDF in a table
  - Use a web calculator
  - Use built-in python functions

STANDARD NORMAL CDF

- The ~~old fashioned~~ classical way to find CDF values is using a table

$$\Phi(1.23) = 0.8907$$

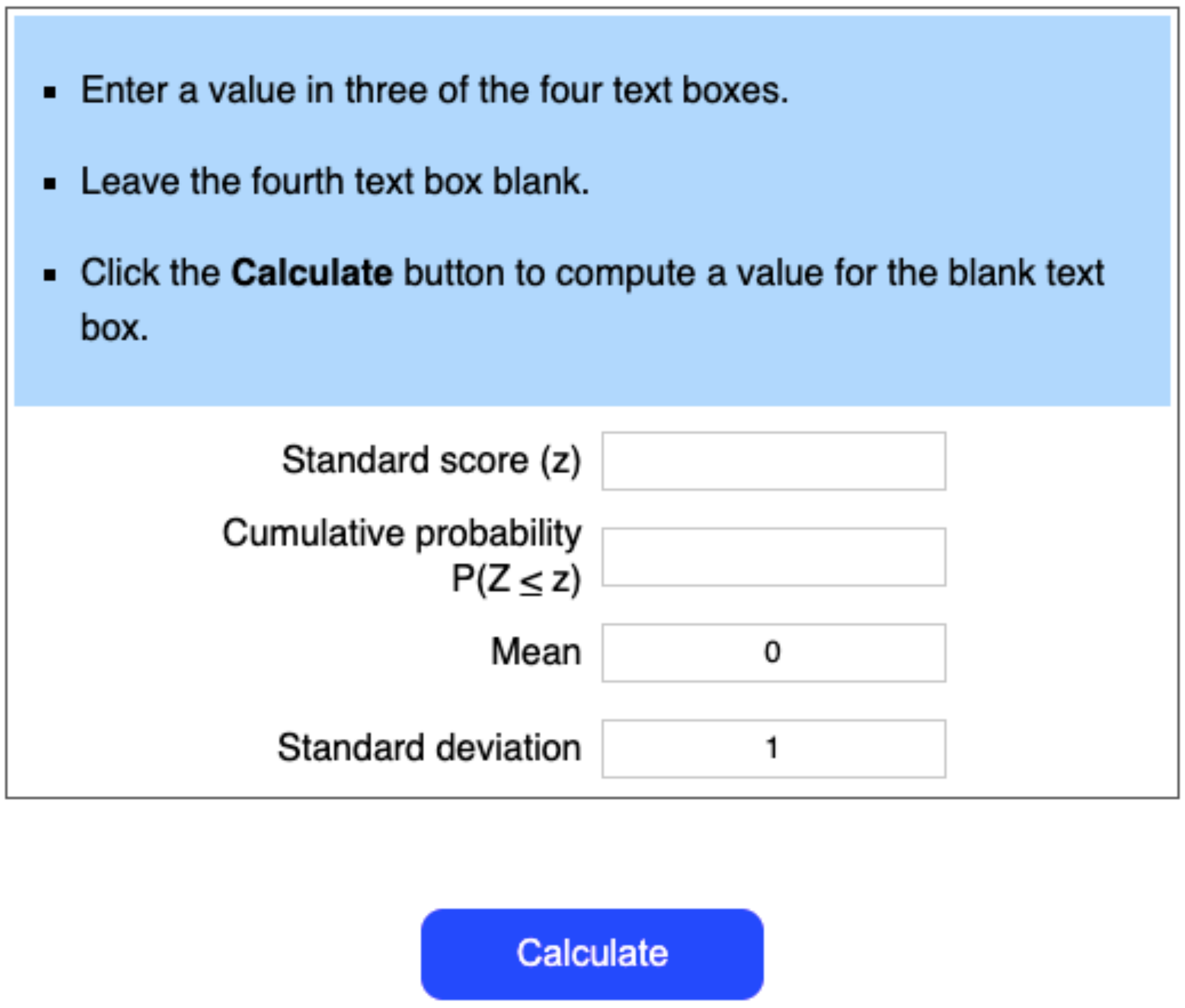
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

STANDARD NORMAL CDF

- The ~~old fashioned~~ classical way to find CDF values is using a table
- Can only look up CDF values for the standard Normal
  - We will see shortly how to use the standard Normal CDF values to evaluate other Normal CDFs
- Can only look up CDF values for non-negative x
  - We will see shortly that we can get the CDF values for negative x from the CDF values from positive x

NORMAL CDF

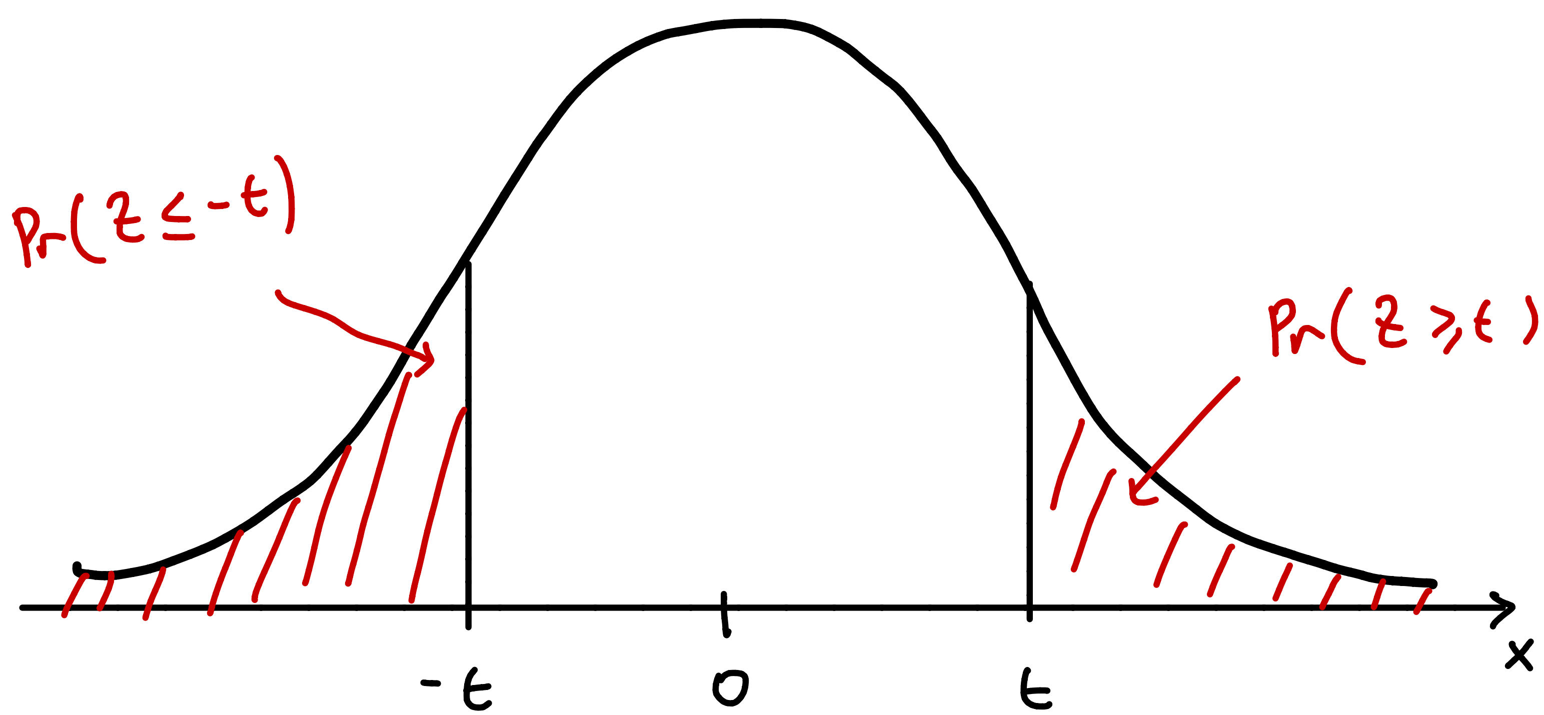
- There are several web calculators, e.g.:
  https://stattrek.com/online-calculator/normal.aspx



NORMAL DISTRIBUTION PROPERTIES: SYMMETRY

- The $\mathrm{Normal}(\mu, \sigma^2)$ PDF is symmetric about $x = \mu$
- In particular, $\mathrm{Normal}(0,1)$ is symmetric about $x = 0$

$$\Pr(Z \leq -t) = \Pr(Z \geq t) \quad \text{for all } t \geq 0$$



NORMAL DISTRIBUTION PROPERTIES: SYMMETRY

- The $\mathrm{Normal}(\mu, \sigma^2)$ PDF is symmetric about $x = \mu$
- In particular, $\mathrm{Normal}(0,1)$ is symmetric about $x = 0$
- The symmetry allows us to get CDF values for negative x from the CDF values for positive x

For all $x \geq 0$, we have:

$$\Phi(-x) = \Pr(Z \leq -x) \underset{\text{symmetry}}{=} \Pr(Z \geq x) = 1 - \Pr(Z \leq x) = 1 - \Phi(x)$$

BOSTON UNIVERSITY

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Alice wants to send a single bit $b \in \{-1, +1\}$ to Bob
- The communication channel is noisy and corrupts the bit by adding random noise $\epsilon \sim \mathrm{Normal}(0,1)$

BOSTON UNIVERSITY

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Alice wants to send a single bit $b \in \{-1, +1\}$ to Bob
- The communication channel is noisy and corrupts the bit by adding random noise $\epsilon \sim \text{Normal}(0,1)$
  - **Example:** Alice sends $b = -1$ and the channel adds noise $\epsilon = 1.23$
  - Bob receives $b + \epsilon = -1 + 1.23 = 0.23$

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Alice wants to send a single bit $b \in \{-1, +1\}$ to Bob
- The communication channel is noisy and corrupts the bit by adding random noise $\epsilon \sim \text{Normal}(0,1)$
- Bob interprets the message as $\hat{b} = \text{sign}(b + \epsilon) \cdot 1$

BOSTON UNIVERSITY

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Alice wants to send a single bit $b \in \{-1, +1\}$ to Bob
- The communication channel is noisy and corrupts the bit by adding random noise $\epsilon \sim \mathrm{Normal}(0,1)$
- Bob interprets the message as $\hat{b} = \mathrm{sign}(b + \epsilon) \cdot 1$
  - **Example:** Alice sends $b = -1$ and the channel adds noise $\epsilon = 1.23$
  - Bob receives $b + \epsilon = -1 + 1.23 = 0.23$
  - Bob received a positive number, he interprets the message as $\hat{b} = 1$
  - In this example, Bob gets the wrong message: $\hat{b} \neq b$

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Alice wants to send a single bit $b \in \{-1, +1\}$ to Bob
- The communication channel is noisy and corrupts the bit by adding random noise $\epsilon \sim \mathrm{Normal}(0,1)$
- Bob interprets the message as $\hat{b} = \mathrm{sign}(b + \epsilon) \cdot 1$
- What is the Pr that Bob gets the wrong message?

Alice sends | Bob receives | Bob decodes

$$\boxed{b} \longrightarrow \boxed{b + \epsilon}$$

$\hat{b} = -1$ if $b + \epsilon < 0$

$\hat{b} = +1$ if $b + \epsilon \geq 0$

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

Alice sends: b → Bob receives: $b + \epsilon$ → Bob decodes:

$$\hat{b} = -1 \text{ if } b + \epsilon < 0$$

$$\hat{b} = +1 \text{ if } b + \epsilon \geq 0$$

Remarks:

- We do not know how Alice chooses which bit to send
- The bit choice does not matter, i.e., the probability that Bob gets the wrong message is the same
- We can see this intuitively via symmetry, and we will argue formally via calculation

BOSTON UNIVERSITY

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Case 1: Alice sends $b = 1$

Alice sends

$$1$$

Bob receives

$$1 + \epsilon$$

Bob decodes

$\hat{b} = -1$ if $1 + \epsilon < 0$

$\hat{b} = +1$ if $1 + \epsilon \geq 0$

$$\Pr(\hat{b} \neq 1) =$$

BOSTON UNIVERSITY

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Case 1: Alice sends $b = 1$

Alice sends: 1 → Bob receives: $1 + \epsilon$ → Bob decodes:

$\hat{b} = -1$ if $1 + \epsilon < 0$

$\hat{b} = +1$ if $1 + \epsilon \geq 0$

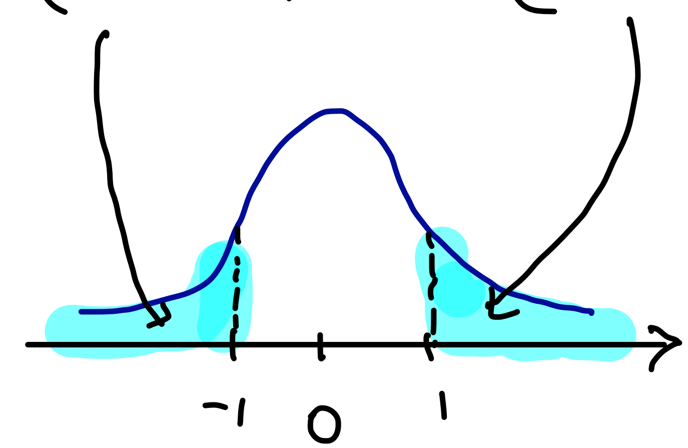
symmetry

$$\Pr(\hat{b} \neq 1) = \Pr(\hat{b} = -1) = \Pr(1 + \varepsilon < 0) = \Pr(\varepsilon < -1) \overset{\text{symmetry}}{=} \Pr(\varepsilon > 1)$$

$$= 1 - \Pr(\varepsilon \leq 1)$$

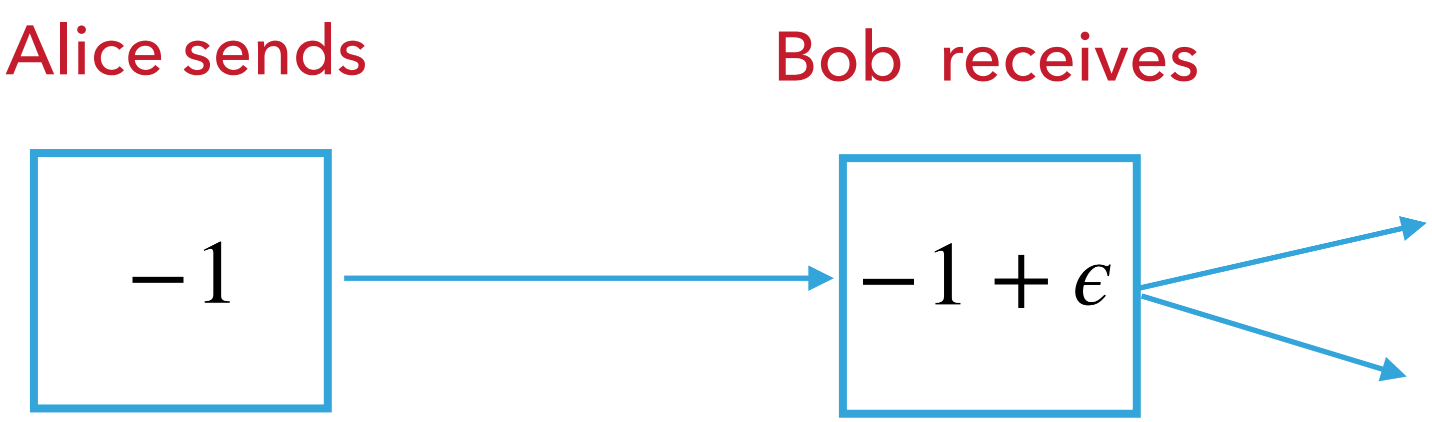
$$= 1 - \Phi(1)$$

$$= 0.1587$$



EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Case 2: Alice sends $b = -1$



Bob decodes

$\hat{b} = -1$ if $-1+\epsilon < 0$

$\hat{b} = +1$ if $-1+\epsilon \geq 0$

$$\Pr(\hat{b} \neq -1) =$$

EXAMPLE: COMMUNICATION ACROSS A NOISY CHANNEL

- Case 2: Alice sends $b = -1$

Alice sends — Bob receives — Bob decodes

$$-1 \longrightarrow -1 + \epsilon$$

$\hat{b} = -1$ if $-1 + \epsilon < 0$

$\hat{b} = +1$ if $-1 + \epsilon \geq 0$

$$\Pr(\hat{b} \neq -1) = \Pr(\hat{b} = +1) = \Pr(-1 + \varepsilon \geq 0) = \Pr(\varepsilon \geq 1)$$

$$= 1 - \Pr(\varepsilon \leq 1) = 1 - \Phi(1) = 0.1587$$

NORMAL DISTRIBUTION PROPERTIES: STANDARDIZATION

- What if we need to look up CDF values for $X \sim \mathrm{Normal}(\mu, \sigma^2)$ but we only have CDF values for the standard $\mathrm{Normal}(0,1)$ distribution?

NORMAL DISTRIBUTION PROPERTIES: STANDARDIZATION

- What if we need to look up CDF values for $X \sim \mathrm{Normal}(\mu, \sigma^2)$ but we only have CDF values for the standard $\mathrm{Normal}(0,1)$ distribution?
- We can do so by "standardizing," as follows:
  - Let $X \sim \mathrm{Normal}(\mu, \sigma^2)$
  - Consider the random variable $Z = \dfrac{X - \mu}{\sigma}$

NORMAL DISTRIBUTION PROPERTIES: STANDARDIZATION

- What if we need to look up CDF values for $X \sim \text{Normal}(\mu, \sigma^2)$ but we only have CDF values for the standard $\text{Normal}(0,1)$ distribution?
- We can do so by "standardizing," as follows:
  - Let $X \sim \text{Normal}(\mu, \sigma^2)$
  - Consider the random variable $Z = \dfrac{X - \mu}{\sigma}$
  - We can show the following:

$$Z \sim \text{Normal}(0,1)$$

BOSTON UNIVERSITY

NORMAL DISTRIBUTION PROPERTIES: STANDARDIZATION

- What if we need to look up CDF values for $X \sim \text{Normal}(\mu, \sigma^2)$ but we only have CDF values for the standard $\text{Normal}(0,1)$ distribution?
- We can do so by "standardizing," as follows:
  - Let $X \sim \text{Normal}(\mu, \sigma^2)$
  - Consider the random variable $Z = \dfrac{X - \mu}{\sigma}$
  - We can show the following:

$$Z \sim \text{Normal}(0,1) \qquad\qquad F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right)$$

BOSTON UNIVERSITY

ASK THE AUDIENCE

Optional example, not covered in lecture

- The annual snowfall in Boston is modeled as a Normal random variable with $\mu = 60$ inches and $\sigma = 20$ inches
- What is the probability that next year's snowfall will be at least 80 inches?

X = next year's snowfall $\quad X \sim \text{Normal}(60, 20^2)$

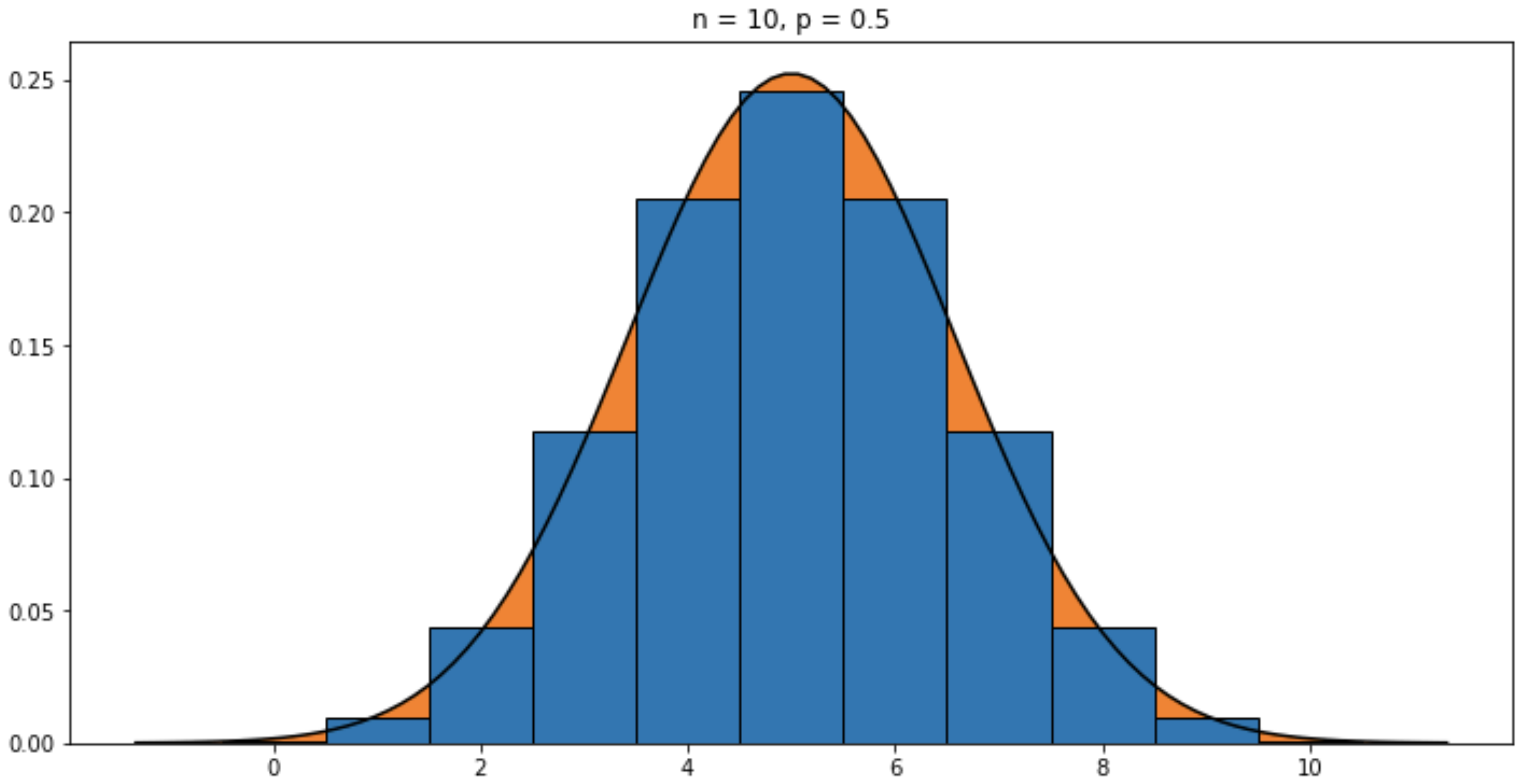
$$\Pr(X \geq 80) = 1 - \Pr(X \leq 80)$$
$$= 1 - \Pr\left(Z \leq \frac{80-60}{20}\right) \quad \text{where } Z = \frac{X-\mu}{\sigma} \sim \text{Normal}(0,1)$$
$$= 1 - \Pr(Z \leq 1) = 1 - \Phi(1) = 0.1587$$

BINOMIAL AND NORMAL

- The Normal distribution can be viewed as a continuous approximation of the Binomial distribution
- We will illustrate this relationship next, by plotting the Binomial and Normal PDFs with the same expectation and variance
    - **Recall:** The Binomial(n, p) distribution has expectation np and variance $np(1-p)$

BINOMIAL AND NORMAL

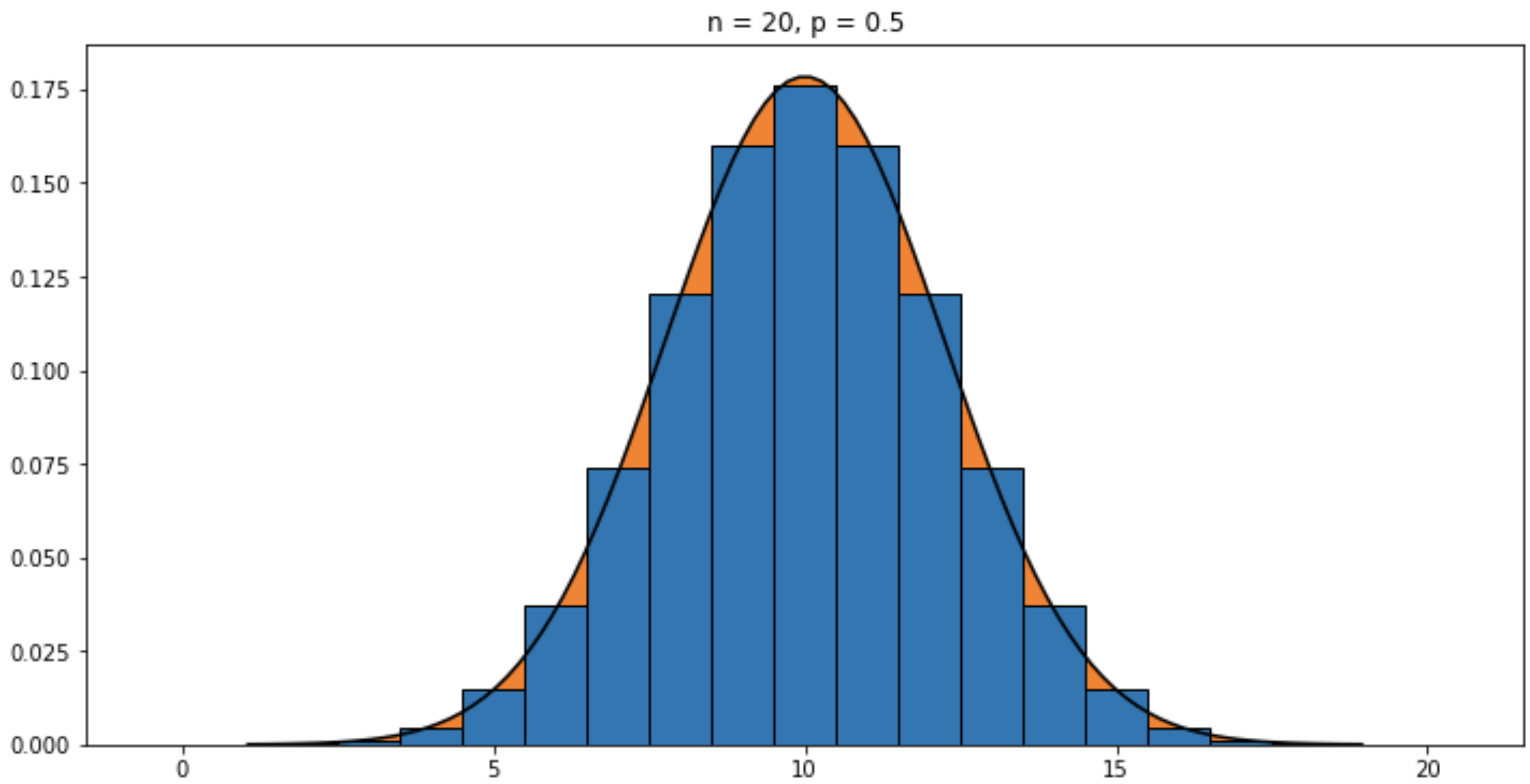
Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



BOSTON UNIVERSITY

BINOMIAL AND NORMAL

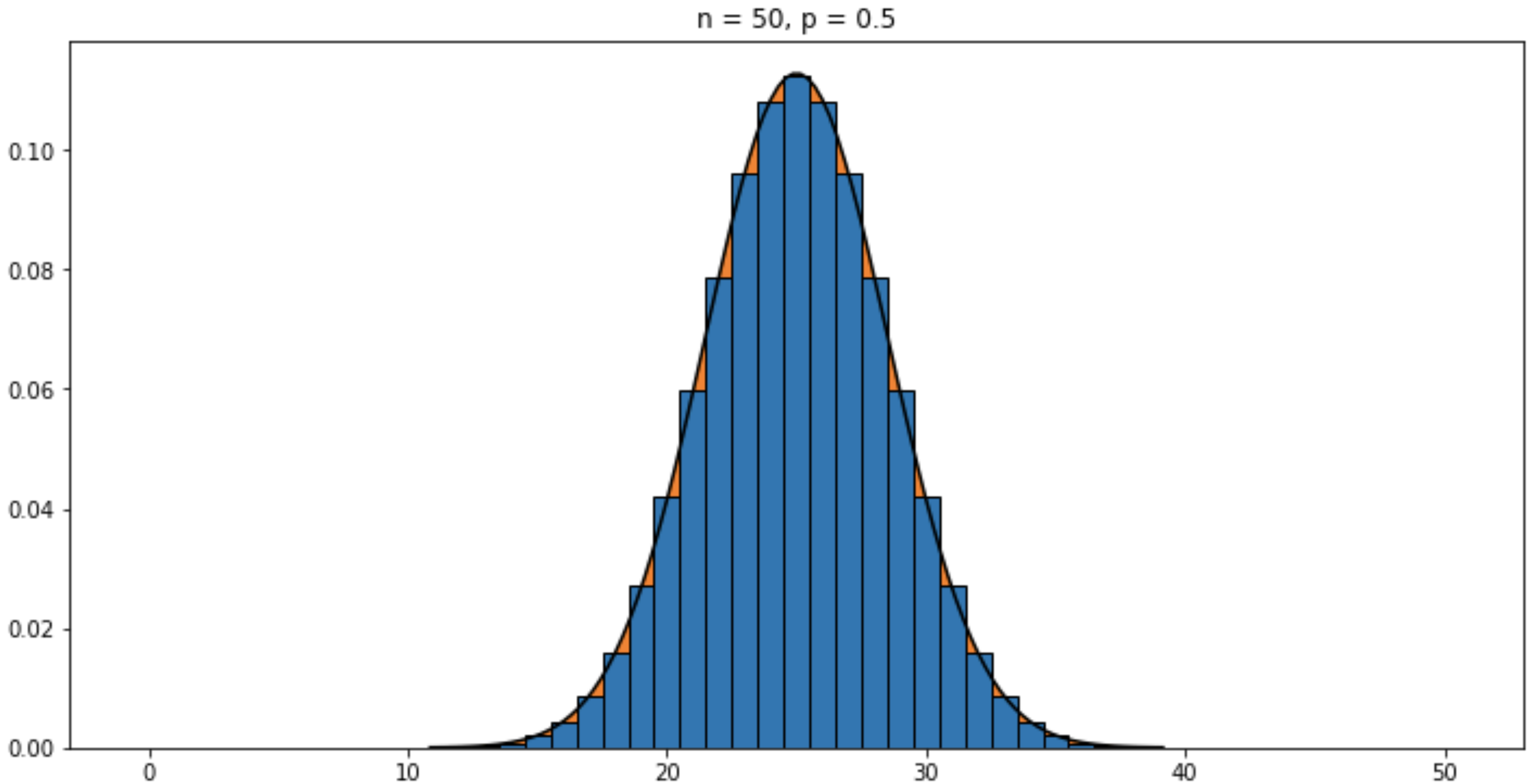
Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



BOSTON UNIVERSITY

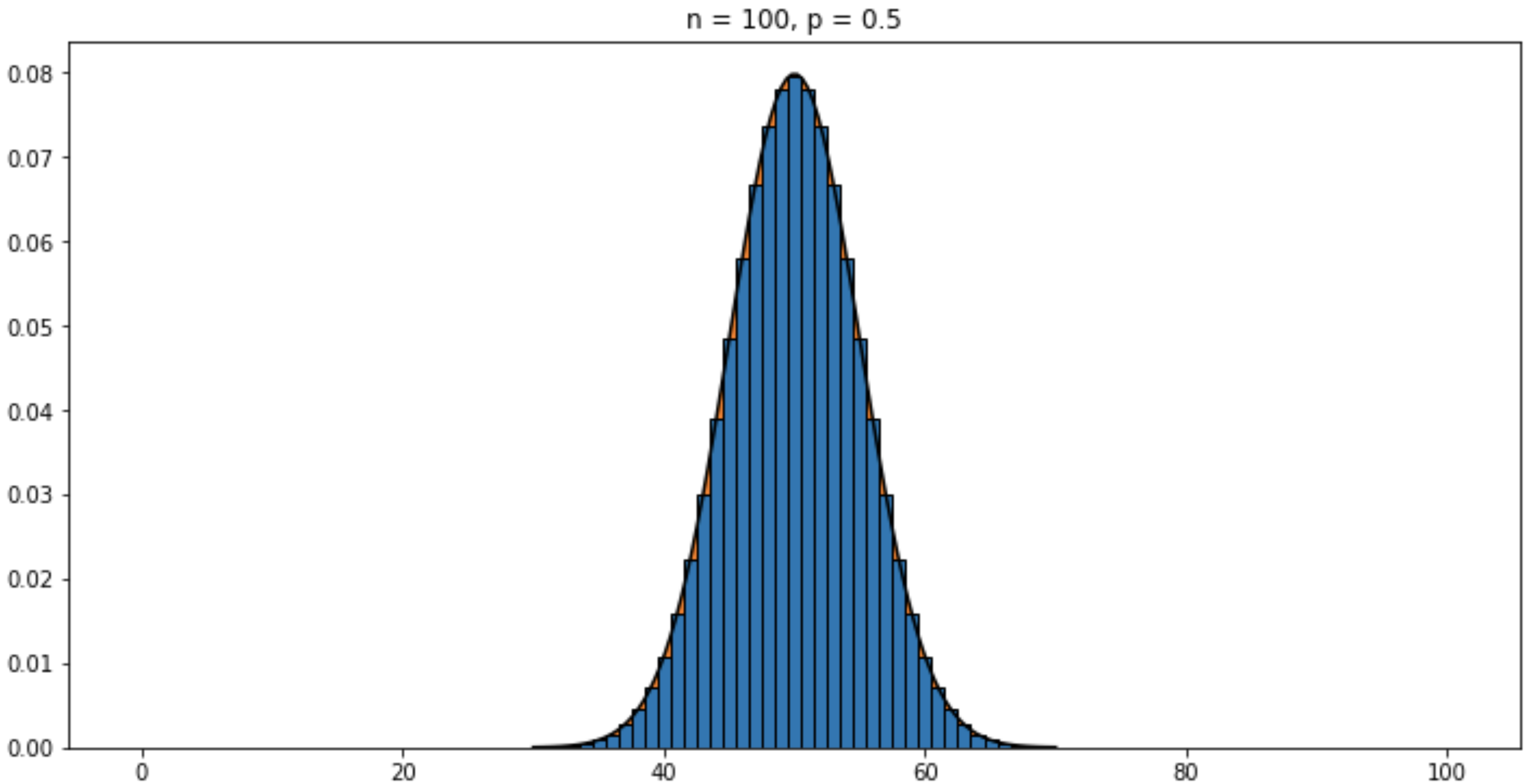
BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



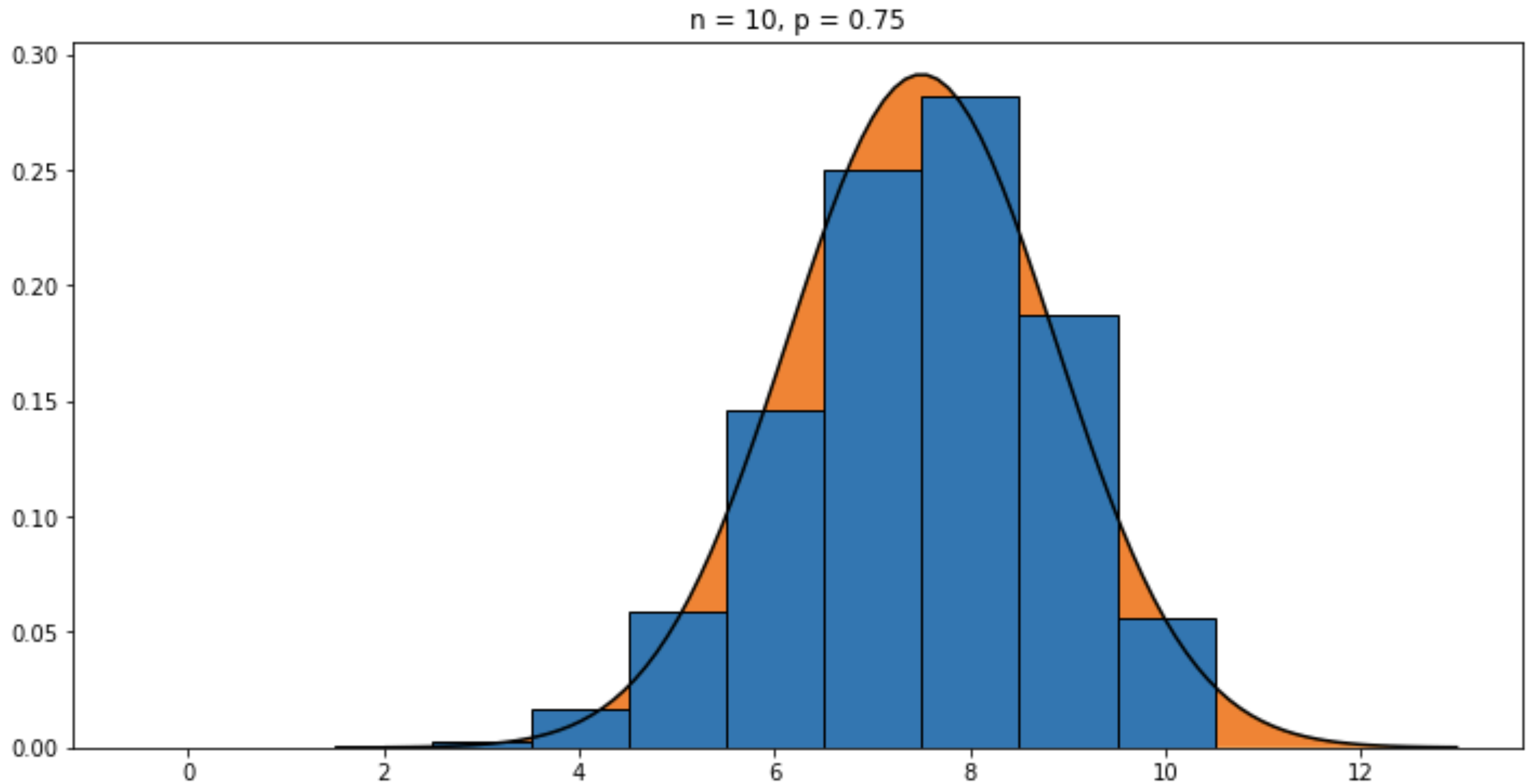
BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



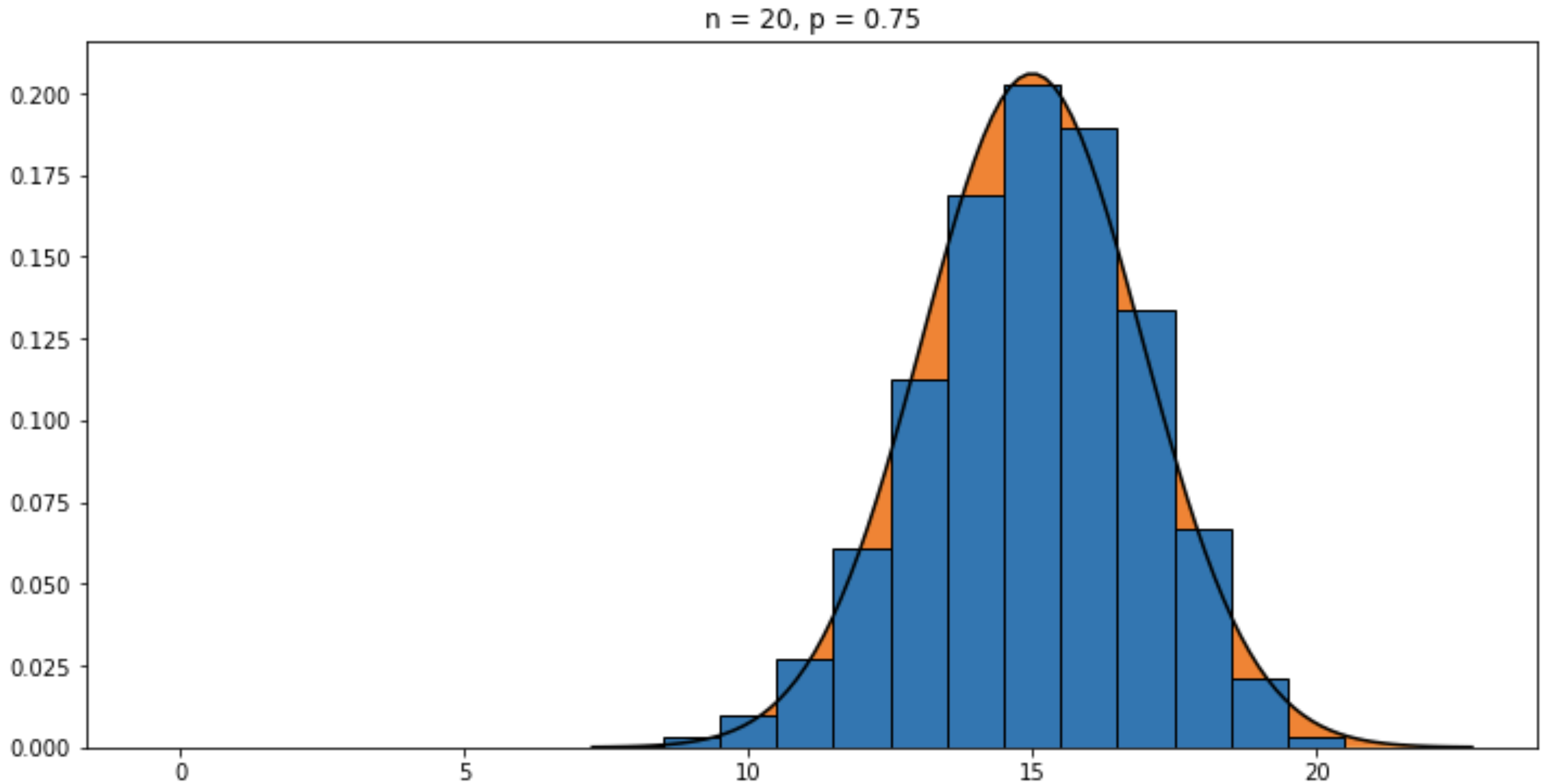
BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



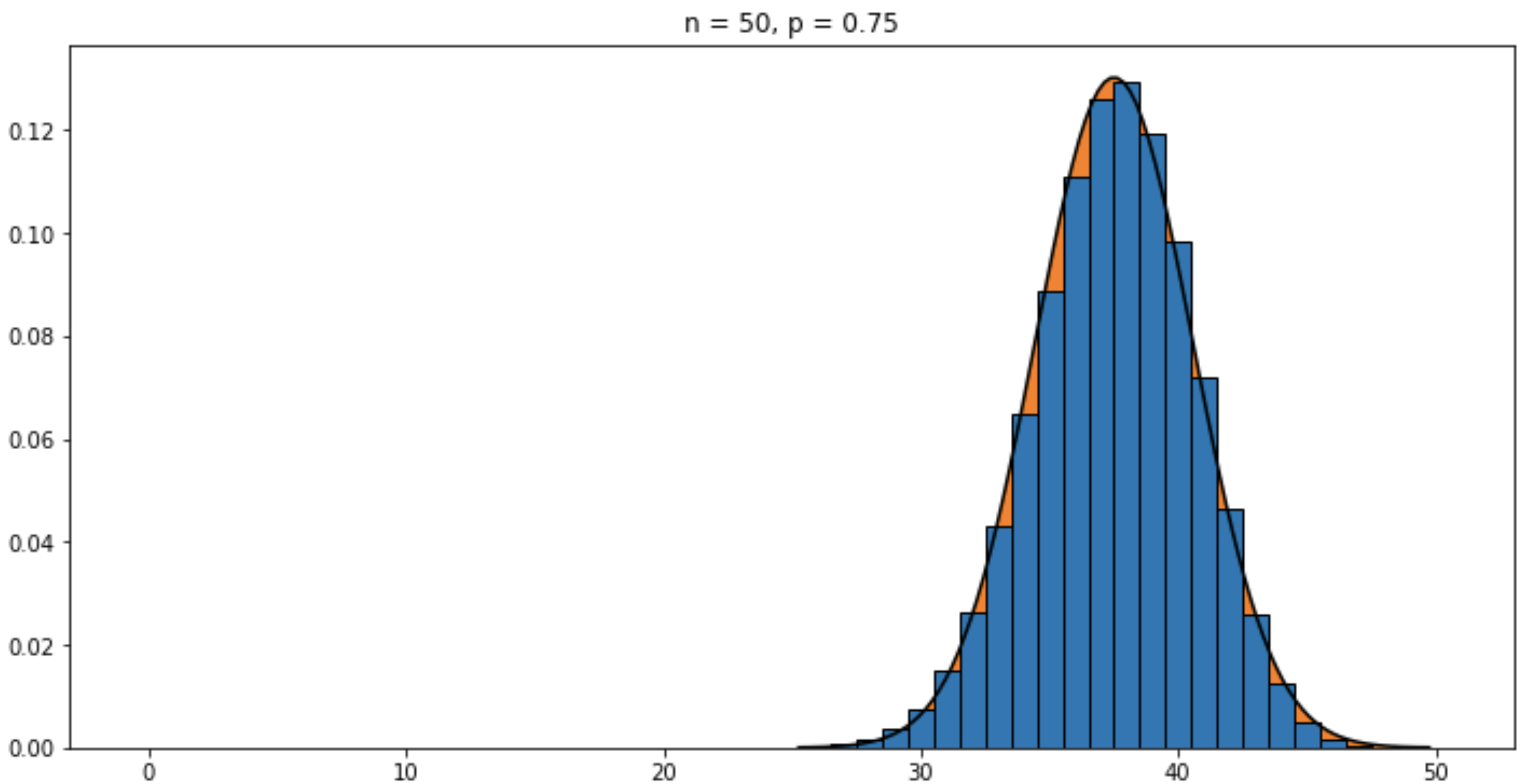
BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))



BINOMIAL AND NORMAL

Plot of Binomial and Normal distributions with the same Ex and Var: Binomial(n, p) and Normal(np, np(1-p))