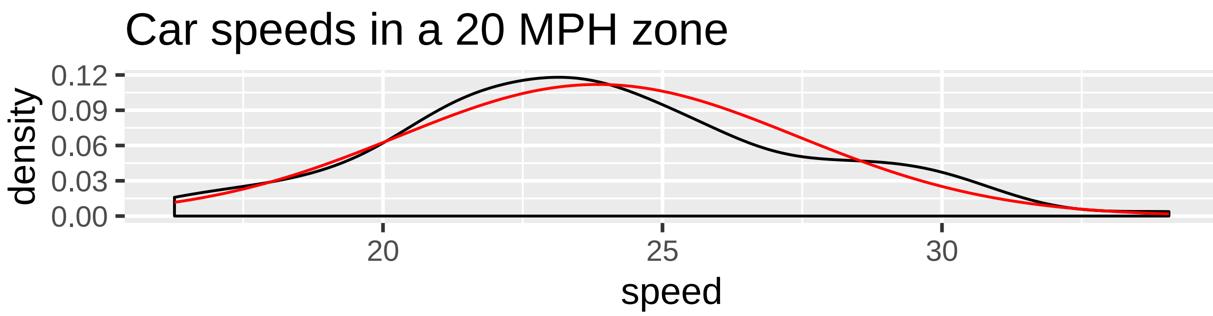
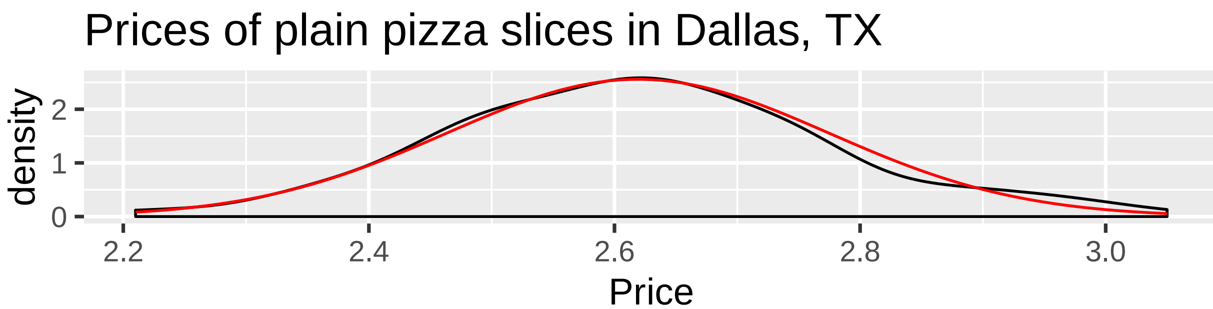
The Normal Distribution (Chapter 5)

Two examples (black: observed sample data; red: a normal model with the same mean and standard deviation)



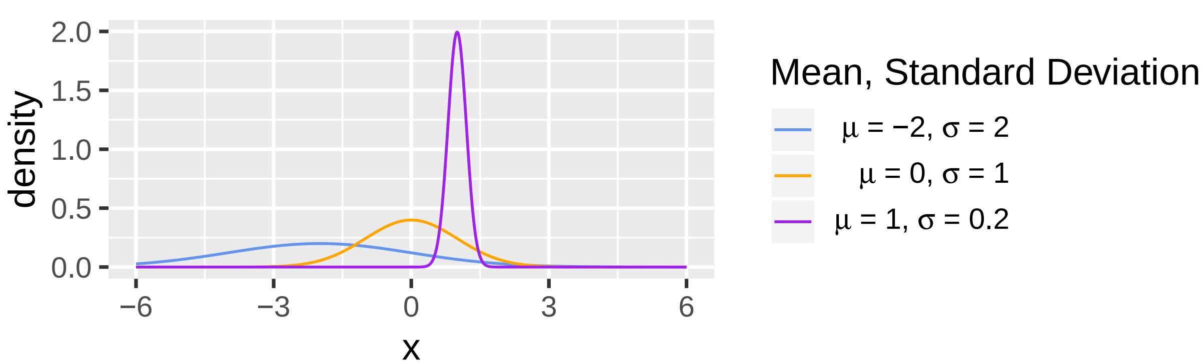


Let the random variable X be the numeric value of one of these variables for a randomly sampled item from the population

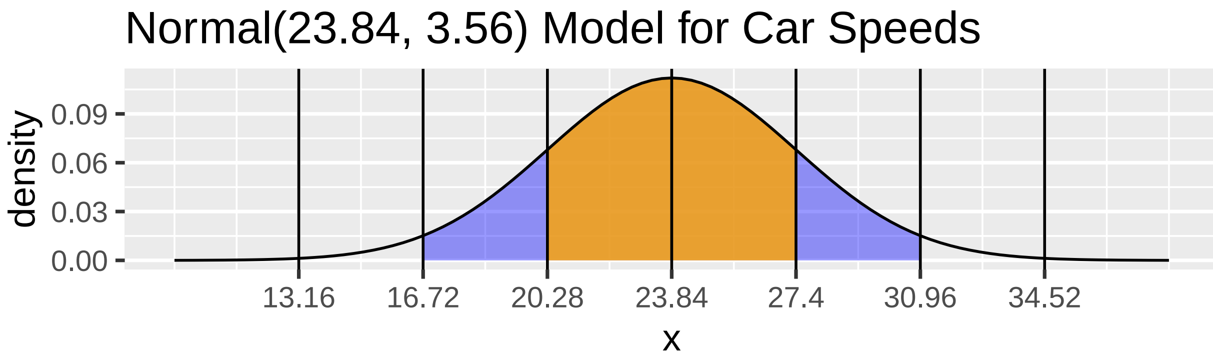
- Example: X is the speed of a randomly selected car in a 20 MPH speed zone.
- Example: X is the price of a piece of cheese pizza from a randomly selected restaurant in Dallas, TX.

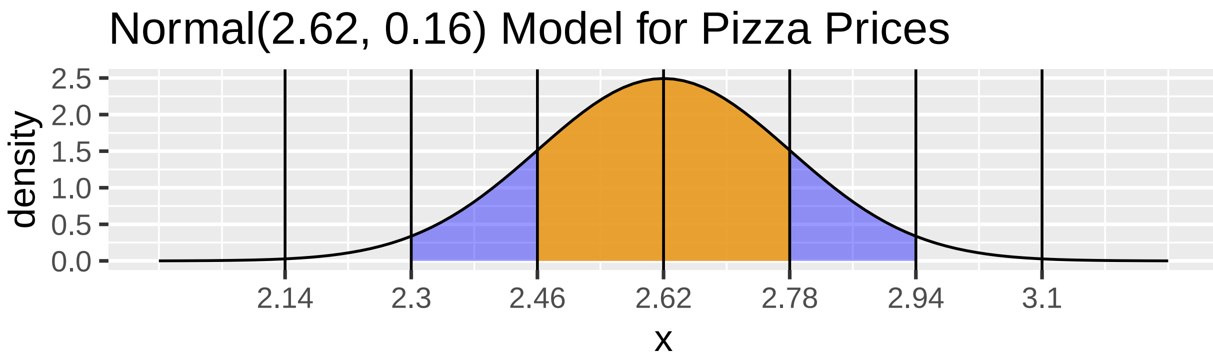
We could **model** the value of X as being a draw from a normal distribution

- $X \sim \text{Normal}(\mu, \sigma)$
- Read: X follows a normal distribution with mean μ ("mew", like a cat) and standard deviation σ ("sigma")
- μ and σ determine the center and spread of the distribution



Here are pictures of normal models for car speeds and pizza prices:





- Area under the curve is the probability of getting an observation in that region
- For any normal distribution,
  - 68% of observations are within $\mu \pm \sigma$ (orange area is about 0.68)
  - 95% of observations are within $\mu \pm 2\sigma$ (sum of orange and blue areas is about 0.95)
  - Total area under the curve is 1 (all observations have some value of x).

z-scores

For calculating probabilities, what matters is **how many standard deviations away from the mean a particular number x is**.

This is the z-score of x: $z = \frac{x-\mu}{\sigma}$

Note: If $X \sim \text{Normal}(\mu, \sigma)$, then $Z \sim \text{Normal}(0, 1)$

Example: Suppose X is the speed of a randomly selected car in a 20MPH zone, and the distribution of speeds of such cars is Normal$(23.84, 3.56)$. How many standard deviations above or below the mean is a car driving 27.4 miles per hour?

Probability calculations in R

The `pnorm` function calculates probabilities involving the normal distribution.

We will provide the z-score as an argument to `pnorm`.

Example: What is the probability that a randomly selected car will be driving less than 27.4 miles per hour, assuming that car speeds follow a Normal$(23.84, 3.56)$ **distribution?**

The number I provided to `pnorm` below is the z-score of 27.4.

```
pnorm(1)
```

```
## [1] 0.8413447
```

Example: What is the probability that a randomly selected car will be driving more than 27.4 miles per hour, assuming that car speeds follow a Normal$(23.84, 3.56)$ **distribution?**

The number I provided to `pnorm` below is the z-score of 27.4.

```
pnorm(1, lower.tail = FALSE)
```

```
## [1] 0.1586553
```

Example: What is the probability that a randomly selected car will be driving exactly 27.4 miles per hour, assuming that car speeds follow a Normal$(23.84, 3.56)$ **distribution?**

There is space to do calculations for the following examples in Lab 12 on Gryd.

Example: Suppose that the price of a slice of cheese pizza from a randomly selected restaurant in Dallas, TX follows a Normal$(2.62, 0.16)$ **distribution. Find the probability that a slice of pizza from a randomly selected restaurant costs less than $2.25.**

Example: Suppose that the price of a slice of cheese pizza from a randomly selected restaurant in Dallas, TX follows a Normal$(2.62, 0.16)$ **distribution. Find the probability that a slice of pizza from a randomly selected restaurant costs more than $3.00.**