Section 4.5: The Normal Distribution

1 The Normal Distribution and the 68-95-99.7 Rule

The **normal curve** (also known as the bell curve) is the most common continuous probability distribution. The normal curve is completely determined by two parameters: mean and standard deviation. The normal curve is symmetric about the mean which is also the median and the mode. Most data is clumped in close to the mean.

The normal distribution is important since it tells us the amount of data that falls in particular intervals relative to the mean and standard deviation. The notation $N(\mu, \sigma)$ states that the data has a normal distribution with mean μ and a standard deviation σ. We use the notation mean μ and standard deviation σ to indicate that these are defining parameters for a statistical distribution rather than statistical values we computed from a sample. For example, $N(35, 2.3)$ indicates a normal distribution with a mean of 35 and a standard deviation of 2.3.

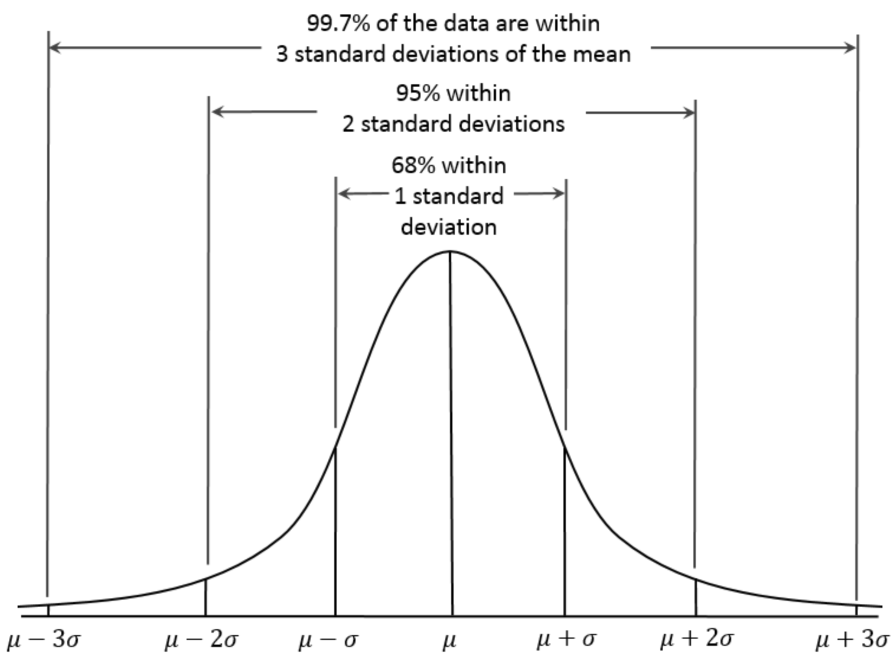
Remark 1 *The Navidi text uses the **very nonstandard** format* $N(\mu, \sigma^2)$.

Remark 2

1.1 The Empirical Rule (68-95-99.7 Rule)

For data distributions that have a **bell-shape distribution** (**normal curve**), the mean and standard deviation tell us a lot about the spread of data from the center.

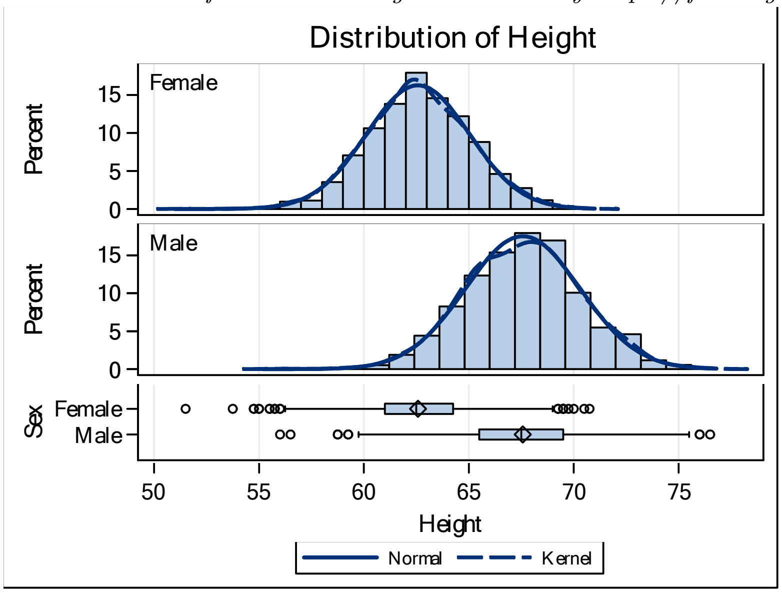
Theorem 3 *The Empirical Rule (68-95-99.7 Rule) states that every normal distribution 68% of the data falls within one standard deviation of the mean, 95% of the data falls within two standard deviations of the mean and 99.7% of the data (almost all) falls within three standard deviations of the mean.*

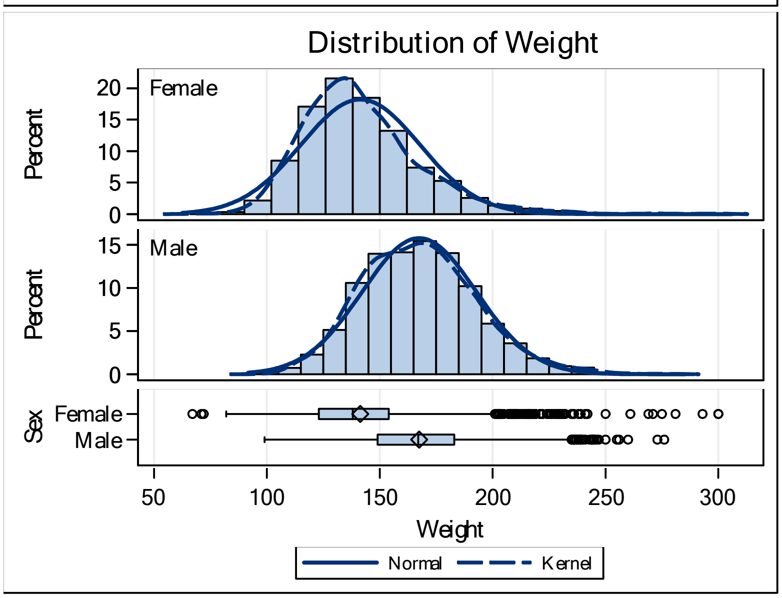


https://en.wikipedia.org/wiki/Normal_distribution

Remark 4 *We frequently denote a normal distribution by* $N(\mu, \sigma)$.

Example 1 *Heights and weights of men and women follow a normal distribution. Data taken from the Framington Heart Study–https://framinghamheartstudy.org/).*





Problem 1 *Heights of men follow a normal distribution with an average of 69" and a standard deviation of 2.8" (I'm rounding to make the arithmetic easier).*

This indicates that the central 68% of men are from   to   tall.

What percentage of men are between 60.6" and 77.4" tall?

What percentage of men are between 69" and 71.8" tall?

Problem 2 *Consider a class whose test results have a distribution of N(75,6).*

What grades comprise the central 68% of the students?

What percentage of grades are between 75 and 81?

What percentage of grades are between 81 and 87?

What percentage of grades are below 57?

If 2000 students took this test, how many students earned a grade less than 57?

The pdf for the normal curve is

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

Finding area under the normal curve is not as simple as when when working with triangles, rectangles and circles. The probability that an observation falls in the interval from a to b in the normal curve is $\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$.

Without resorting to integration, a useful tool for computing probabilities in the normal curve is to convert to standardized units is the z-score

$$z = \frac{x-\mu}{\sigma}.$$

When using a table from the book to compute normal curve values, z-scores are needed. This is not the case when using technology.

2 The Normal Curve and the TI 83/84

The 68-95-99.7 rule is a nice beginning from which to explore the normal curve. It would be greatly limiting if we could only work with z-scores of $\pm 1, \pm 2$ and ± 3. Fortunately, technology allows us to work with any z-score in any normal distribution.

The TI 83/84 series of calculators has two basic types of normal curve commands;

1. Find the percentage p of data between two values z_1 (lower bound) and z_2 (upper bound) in the normal distribution defined by μ and σ:
   normalcdf(z_1 **(lower bound)**, z_2 **(upper bound)**, μ, σ)

2. Find the value z such that p percent of data falls below (to the left of) z in the normal distribution defined by μ and σ:
   invNorm(p, μ, σ)**.** Recall that this is the p^{th} percentile.

Definition 5 *The **standard normal curve** is the normal distribution where* $\mu = 0$ *and* $\sigma = 1$*.*

Problem 3 *What is the probability of an observation between 0 and 1.34 in the standard normal curve?*

Problem 4 *What is the probability of an observation less than* $z = 1.34$ *in the standard normal curve?*

Problem 5 *Find the probability that an observation falls between 0 and .78 in the standard normal curve?*

Problem 6 *In the standard normal curve, find* P_{30}.

Problem 7 *In the standard normal curve, find* P_{78}.

Problem 8 *Find the probability that an observation falls above -1.62 in a normal distribution where the mean is 2 and the standard deviation is 3.*

Problem 9 *Consider a normal distribution where the mean is 56.8 and the standard deviation is 5.5. What is the probability of an observation between 60 and 70?*

Problem 10 *Consider a normal distribution where the mean is 10 and the standard deviation is 15. What is the probability of an observation larger than 20?*

Problem 11 *Consider a normal distribution where the mean is 55 and the standard deviation is 15. What is the 63rd percentile?*

Problem 12 *Consider a normal distribution where the mean is 178.2 and the standard deviation is 15.8. What is the 23rd percentile?*

3 Applications of the Normal Distribution

Problem 13 *Scores on a particular test follow a normal distribution with a mean of 75.6 and a standard deviation of 7.8.*
i. What percentage of students scored above a 90?

ii. Find the score separating the top 15% of the scores from the bottom 85% of scores.

iii. Between what two values do the central 40% of scores fall?

Problem 14 *On a particular track team mile running times follow a normal distribution with a mean of 8.5 minutes and a standard deviation of 1.2 minutes.*
i. What percentage of the track team runs a mile in under 7 minutes?

ii. What percentage of the track team runs a mile in between 8 and 9.3 minutes?

iii. What time separates the fastest 10% of the runners from the rest of the team?

iv. The slowest 20% of all runners will be cut from the team. How fast of a mile does one need to run in order to stay on the team?

Problem 15 *Scores on an exam were normally distributed where 10% of the scores fell below 50 and 5% of the scores fell above 92. Find the mean and standard deviation.*
Note that the information has been given to you in the form of percentile scores: $P_{10} = 50$ *and* $P_{95} = 92$*. In the standard normal curve,* $P_{10} = -1.28$ *and* $P_{95} = 1.645$*. Thus,*

$$\begin{aligned} -1.28 &= \frac{50 - \mu}{\sigma} \\ &\rightarrow -1.28\sigma = 50 - \mu \\ &\rightarrow \mu = 50 + 1.28\sigma. \end{aligned}$$

Plugging this into our other expression shows

$$\begin{aligned} 1.645 &= \frac{92 - \mu}{\sigma} \\ &\rightarrow 1.645\sigma = 92 - \mu \\ &\rightarrow \mu = 92 - 1.645\sigma. \end{aligned}$$

Thus,

$$\begin{aligned} 92 - 1.645\sigma &= \mu = 50 + 1.28\sigma \\ &\rightarrow 92 - 50 = 1.645\sigma + 1.28\sigma \\ &\rightarrow 42 = 2.925\sigma \\ &\rightarrow 14.36 = \sigma. \end{aligned}$$

And now,

$$\begin{aligned} \mu &= 50 + 1.28\sigma \\ &= 50 + 1.28 * 14.36 \\ &= 68.381 \end{aligned}$$

4 Exercises

1. Navidi Section 4.5: 1-8, 10, 12, 15, 19