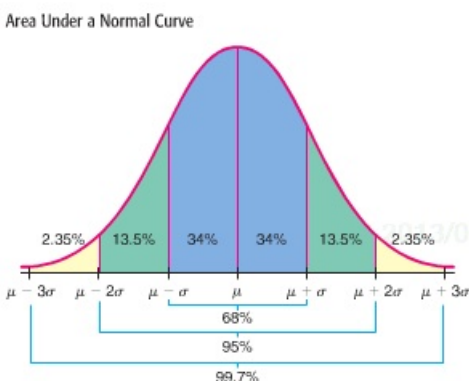
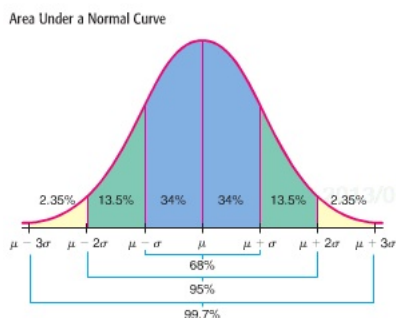


1. Given a normal curve (shown here), what percentage of the area lies



- (a) to the right of μ ?
The mean μ sits in the center of the data, so 50% lies to the right (and to the left) of μ .
- (b) between $\mu - \sigma$ and μ ?
The first region to the left of μ to $\mu - \sigma$ there is 34% of the area.
- (c) between $\mu + \sigma$ and $\mu + 2\sigma$?
Looking at the graph, the region between $\mu + \sigma$ and $\mu + 2\sigma$ consists of 13.5% of the area.
- (d) between $\mu - 2\sigma$ and $\mu + \sigma$?
For $\mu - 2\sigma$ to $\mu + \sigma$ we can add the three percentages 13.5%, 34%, and 34% to get 81.5%.
2. The incubation time for Rhode Island Red chicks is normally distributed with a mean of $\mu = 21$ days and a standard deviation of about $\sigma = 1$ day. If 1000 eggs are being incubated, how many chicks do we expect to hatch
- (a) on day 21 or after?
Since $\mu = 21$, the percentage of the eggs that will hatch at that time or greater is 50%. So, that is 500 of the eggs.
- (b) between 20 and 21 days?
Since $\mu = 21$ and $\sigma = 1$, we are looking at the percentage of the eggs in the interval $\mu - \sigma$ to μ which is 34%. So, that is 340 of the eggs.
- (c) between 22 and 23 days?
Since $\mu = 21$ and $\sigma = 1$, we are looking at the percentage of eggs in the interval $\mu + \sigma$ and $\mu + 2\sigma$ which is 13.5%. So, that is 135 of the eggs.
- (d) between 19 and 22 days?
Since $\mu = 21$ and $\sigma = 1$, we are looking at the percentage of eggs in the interval $\mu - 2\sigma$ and $\mu + \sigma$ which is 81.5%. So, that is 815 of the eggs.

3. The Empirical Rule tells us the percentage of data values that lie within certain intervals about the mean when we have a normal distribution as follows.



- (a) According to Chebyshev's Theorem, what is the minimum percentage of data that must lie within the interval $\mu - 2\sigma$ to $\mu + 2\sigma$? How does this compare to the results for the Empirical Rule?

The result of Chebyshev's Theorem is that at least $1 - \frac{1}{2^2} = 75\%$ of the data fall within in the interval...and this is true for *any* data distribution.

For a normal distribution, the Empirical Rule shows us that there is approximate 95% of the data falling within that same range.

- (b) The diagram for the Empirical Rule only accounts for 99.7% of the data values. Where do the remaining 0.3% of the data fall?

The remaining 0.3% must lie outside the interval from $\mu - 3\sigma$ to $\mu + 3\sigma$...That is, those data values lie fall out in the tails of the distribution. Additionally, since the distribution is symmetric, half of them (or 0.15%) lie below $\mu - 3\sigma$ and the other half lie above $\mu + 3\sigma$.

- (c) Keeping in mind that remaining 0.3% of the data, what percentage of the data values lie to the left of $\mu - \sigma$?

Add 13.5%, 2.35%, and 0.15% to get 16%.

Note that equivalently, since we know that 50% lie to the left of the mean μ and 34% lie between $\mu - \sigma$ and μ , that leaves 16% lying to the left of $\mu - \sigma$.

- (d) Keeping in mind that remaining 0.3% of the data, what percentage of the data values lie to the left of $\mu + 2\sigma$?

Add 13.5%, 34%, 34%, 13.5%, 2.35%, and 0.15% to get 97.5%.

Note that equivalently, since we know that 50% lie to the left of the mean μ and 34% + 13.5% lie between μ and $\mu + 2\sigma$, that total is 97.5% lying to the left of $\mu + 2\sigma$.

4. In preparation for Sections 6.2 and 6.3 (of Brase and Brase), suppose the mean of a normal distribution is $\mu = 75$ with a standard deviation of $\sigma = 8$. For each of the following values of the random variable x , determine whether x is above or below the mean and then determine the number of standard deviations away from the mean that x lies. (Note that your second answer could be a fraction or decimal number.)

(a) $x = 83$

$x = 83$ is above the mean of $\mu = 75$. Specifically, $\frac{83-75}{8} = 1$. So, x is 1 standard deviation above the mean.

(b) $x = 51$

$x = 51$ is below the mean of $\mu = 75$. Specifically, $\frac{51-75}{8} = -3$. So, x is 3 standard deviations below the mean.

(c) $x = 79$

$x = 79$ is above the mean of $\mu = 75$. Specifically, $\frac{79-75}{8} = 0.5$. So, x is 0.5 standard deviations above the mean.

(d) $x = 65$

$x = 65$ is below the mean of $\mu = 75$. Specifically, $\frac{65-75}{8} = -1.25$. So, x is 1.25 standard deviations below the mean.

5. Hydraulic pressure in the main cylinder of the landing gear of a commercial jet is very important for a safe landing. In-flight landing tests show that the actual pressure in the main cylinders is variable (approximately normal) with a mean of 819 pounds per square inch and a standard deviation of 23 pounds per square inch, which are considered safe.

Two different planes were tested with 10 test-landings. Here are the recorded pressures for each.

Plane A

| Landing number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pressure | 870 | 855 | 830 | 815 | 847 | 836 | 825 | 810 | 792 | 820 |

Plane B

| Landing number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pressure | 865 | 850 | 841 | 820 | 815 | 789 | 801 | 765 | 730 | 725 |

- (a) Is the pressure for Plane A “in control” or not? If not, describe the specific out-of-control signals present.

Out of control Signal 1: $\mu - 3\sigma = 750$ and $\mu + 3\sigma = 888$. No values fall outside that range.

Out of control Signal 2 : $\mu = 819$. There are no more than 3 consecutive values above μ and no more than 2 consecutive values below μ .

Out of control Signal 3: $\mu - 2\sigma = 773$ and $\mu + 2\sigma = 865$. There is only 1 value that lies beyond either end of that range, so certainly not 2 of 3 consecutive values do. There are no out-of-control warning signals present for Plane A.

- (b) Is the pressure for Plane B “in control” or not? If not, describe the specific out-of-control signals present.

Out of control Signal 1: $\mu - 3\sigma = 750$ and $\mu + 3\sigma = 888$. There are two values that fall below 750.

Out of control Signal 2 : $\mu = 819$. There are no more than 4 consecutive values above μ and no more than 6 consecutive values below μ .

Out of control Signal 3: $\mu - 2\sigma = 773$ and $\mu + 2\sigma = 865$. The last three values recorded lie below 773, so there are two occasions of 2 out of 3 consecutive values flying beyond the 2σ level on the same side. There are two out-of-control warning signals present for Plane B.

6. The college physical education department offered an advanced first aid course last semester. The scores on the comprehensive final exam were normally distributed, and the z scores for some of the students are shown below:

{Robert, 1.10} {Juan, 1.70} {Susan, -2.00} {Joel, 0.00} {Jan, -0.80} {Linda, 1.60}

If the mean score was $\mu = 150$ with standard deviation $\sigma = 20$, what was the final exam score for each student?

Robert: $z = 1.10$, so the raw score is $1.10 \cdot 20 + 150 = 172$.

Juan: $z = 1.70$, so the raw score is $1.70 \cdot 20 + 150 = 184$.

Susan: $z = -2.00$, so the raw score is $-2.00 \cdot 20 + 150 = 110$.

Joel: $z = 0.00$, so the raw score is $0.00 \cdot 20 + 150 = 150$.

Jan: $z = -0.80$, so the raw score is $-0.80 \cdot 20 + 150 = 134$.

Linda: $z = 1.60$, so the raw score is $1.60 \cdot 20 + 150 = 182$.

7. Let z be a random variable with a *standard* normal distribution. Find the indicated probability.

(a) $P(z \geq 0)$

This is simply 50% because $z = 0$ sits at the mean, and 50% of the area is above the mean.

(b) $P(z \leq 3.20)$

`normalcdf(-1E99, 3.20, 0, 1) = 0.9993127979 \approx 99.93%`

(c) $P(z \geq -1.50)$

`normalcdf(-1.50, 1E99, 0, 1) = 0.9331927713 \approx 93.32%`

(d) $P(-1.78 \leq z \leq 1.40)$

`normalcdf(-1.78, 1.40, 0, 1) = 0.8817053583 \approx 88.17%`

8. Let x be a random variable with a normal distribution with mean $\mu = 15$ and standard deviation $\sigma = 4$. Find the indicated probability.

(a) $P(x \leq 12)$

`normalcdf(-1E99, 12, 15, 4) = 0.2266272794 \approx 22.66%`

(b) $P(x > 10)$

`normalcdf(10, 1E99, 15, 4) = 0.894350161 \approx 89.44%`

(c) $P(15 \leq x \leq 20)$

`normalcdf(15, 20, 15, 4) = 0.3943501605 \approx 39.44%`

(d) $P(8 < x < 22)$

`normalcdf(8, 22, 15, 4) = 0.9198817729 \approx 91.99%`

9. Find the z -value described.

(a) Find z such that 5.2% of the standard normal curve lies to the left of z .

This is a left-tail area, so $z = \text{invNorm}(0.052, 0, 1) = -1.625763385$.

- (b) Find z such that 5% of the standard normal curve lies to the right of z .
Since this is a right-tail area, we can first translate this to a left-tail area by equating “5% to the right of z ” with “95% to the left of z .”
Then $z = \text{invNorm}(0.95, 0, 1) = 1.644853626$.
Note that alternatively, symmetry can be used with $\text{invNorm}(0.05, 0, 1) = -1.644853626$ to conclude $z = 1.644853626$.
- (c) Find z such that 95% of the standard normal curve lies between $-z$ and $+z$.
To use `invNorm`, we must first translate this to a left-tail area problem. If 95% of the area lies *between* $-z$ and $+z$, then we can equate this to 97.5% lies to the left of $+z$. Then $z = \text{invNorm}(0.975, 0, 1) = 1.959963986$, and $-z = -1.959963986$.
10. Suppose that the scores for a Chemistry midterm had a normal distribution with mean $\mu = 71.4$ and a standard deviation of $\sigma = 8.3$.
- (a) Find the exam score for a student that scored in the 80th percentile.
The 80th percentile refers to the score where 80% of the class earned that score or below. That is, the left-tail area related to this score is 80%. So, $x = \text{invNorm}(0.80, 71.4, 8.3) = 78.38545624 \approx 78.4$.
- (b) Find the exam score for a student that scored in the 25th percentile.
The 25th percentile refers to the score where 25% of the class earned that score or below. That is, the left-tail area related to this score is 25%. So, $x = \text{invNorm}(0.25, 71.4, 8.3) = 65.80173508 \approx 65.80$.
- (c) Find the exam score for a student who score lower than 60% of the class.
Scoring lower than 60% of the class implies that the right-tail area related to this score is 60%. However, since the `invNorm` function requires the left-tail area as its input, we must first translate this problem to the left-tail area of 40%. Then, $x = \text{invNorm}(0.40, 71.4, 8.3) = 69.29721906 \approx 69.30$.
11. Suppose the distribution of heights of American men (20 years of age and older) is approximately normal with a mean of 69.4 inches and a standard deviation of 3 inches.
- (a) What is the z value for a height of 6 feet?
First we convert 6 feet to 72 inches. Then $z = \frac{72-69.4}{3} = 0.86666667$.
- (b) What percentage of American men (20 years of age and older) are shorter than 6 feet?
We can use `normalcdf`.
$$P(z < 0.86666667) = \text{normalcdf}(-1E99, 0.86666667, 0, 1) = 0.8069377087 \approx 80.69\%$$
- (c) What is the z value for a height of 6 feet 4 inches?
First we convert 6 feet 4 inches to 76 inches. Then $z = \frac{76-69.4}{3} = 2.2$.
- (d) What percentage of American men (20 years of age and older) are taller than 6 feet 4 inches?
We can use `normalcdf`.
$$P(z > 2.2) = \text{normalcdf}(2.2, 1E99, 0, 1) = 0.0139033989 \approx 1.39\%$$

12. Accrotime is a manufacturer of quartz crystal watches. Accrotime researchers have shown that the watches have an average life of 28 months before certain electronic components deteriorate, causing the watch to become unreliable. The standard deviation of watch lifetimes is 5 months, and the distribution of lifetimes is normal.

- (a) If Accrotime guarantees a full refund on any defective watch for 2 years, what percentage of total production will the company expect to replace?

First we need to make sure all time intervals are measured in the same units. So, we are looking for the percentage of watches that are expected to fail in 24 months or fewer.

$P(x \leq 24) = \text{normalcdf}(-1E99, 24, 28, 5) = 0.2118553337$. This means that the company should expect to replace about 21.19% of its watches.

- (b) If Accrotime does not want to make refunds on more than 12% of the watches it makes, how long should the guarantee period be (to the nearest month)?

Certainly, it is shorter than 24 months, since that led to 21.19% replaced. To determine a more precise number of months, we can use `invNorm` where we are looking for a left-tailed area equal to 0.12.

$\text{invNorm}(0.12, 28, 5) = 22.12506604$. This implies that the guarantee period should be 22 months.

This section of the problem set is an activity with the goal of simulating a sampling distribution of sample means. The first part of the activity is for each of you to generate a list of sample means. After the data is gathered, the second part directs you to use the results of the entire class to analyze the outcome.

PART 1

13. Attached to this page is a copy of a random number table for the population of digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Assuming that this random distribution is uniform:

- (a) What is the population mean of the digits?

Since the frequency of all digits is equally weighed, $\mu = \frac{0+1+2+3+4+5+6+7+8+9}{10} = 4.5$.

- (b) What is the population standard deviation of the digits?

This could be done by hand (like the mean just was) or entering $L_1 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, `1-VarStats` indicates that $\sigma = 2.872281323$.

14. Repeat the following steps 10 times. In the end, you will have generated 10 different sample means. Please do NOT simply copy a classmate's values on this part, but please do your own work so that you really are generating your own list of sample means.

- (a) *Randomly* select a spot on the random number table.
- (b) Write down the next 10 digits listed from that spot horizontally (if you reach the end of a row, wrap around to the start of the next row to complete your list of 10).
- (c) Compute the mean of that sample of 10 random digits.

| Sample | List of 10 digits | Sample mean |
|--------|-------------------|-------------|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |

PART 2

- (a) Looking at the list of sample means, there should be a variety of values, but they are not uniformly distributed. What is the shape of the distribution for the sample means?

This will certainly vary from section to section, but the distributions should be roughly single-mound, symmetric.

- (b) What is the mean of the sample means?

Again, the specific answers will vary from section to section, but the outcome should be close to the population mean of 4.5.

- (c) What is the standard deviation of the sample means?

Again, the specific answers will vary from section to section, but the outcome should not be too far from $\frac{2.872281323}{\sqrt{10}} \approx 0.9082951$.

- (d) In Section 6.5, one of the most critical theorems in statistics is introduced; the Central Limit Theorem. Essentially it says that a sampling distribution for sample means (where the sample size for all samples is n) will have a distribution close to a normal distribution. Further, the mean of the sample means will equal the population mean μ and the standard deviation of the sample means will equal $\frac{\sigma}{\sqrt{n}}$. How did the simulation the class do match up to the theorem?

- i. Was the shape of the distribution of sample means approximately normal?

Answers will vary.

- ii. Was the mean of the sample means close the the population mean μ ?

Answers will vary.

- iii. What was the sample size for our sample means?

$n = 10$

iv. Was the standard deviation of the sample means close to the value $\frac{\sigma}{\sqrt{(n)}}$?

Answers will vary, but $\frac{\sigma}{\sqrt{(n)}} = \frac{2.872281323}{\sqrt{(10)}} \approx 0.9082951$.

| | | | | | | | |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 6 1 9 2 5 | 8 0 2 9 6 | 8 1 0 4 8 | 9 7 1 9 1 | 6 5 1 8 2 | 6 2 8 2 3 | 1 0 7 4 2 | 1 6 8 5 9 |
| 9 4 2 9 5 | 4 0 3 3 9 | 0 4 5 5 3 | 1 9 6 5 6 | 7 7 6 4 5 | 1 5 0 4 4 | 7 2 8 1 5 | 4 2 6 3 8 |
| 2 0 4 8 9 | 3 8 4 5 5 | 9 5 7 2 4 | 7 4 6 4 6 | 7 4 4 4 1 | 9 0 3 9 7 | 4 8 6 3 0 | 5 5 8 4 3 |
| 5 0 1 9 6 | 0 6 3 3 0 | 7 7 5 6 8 | 8 7 0 5 4 | 8 2 8 5 8 | 5 9 6 9 2 | 6 2 1 7 1 | 2 7 8 8 7 |
| 2 8 4 0 1 | 9 7 5 1 0 | 6 3 5 0 7 | 1 4 3 9 6 | 3 9 0 0 5 | 5 7 0 7 8 | 7 1 0 0 2 | 2 4 4 2 5 |
| 7 0 1 6 6 | 5 1 4 2 1 | 8 4 0 5 5 | 2 5 1 4 6 | 8 1 4 7 6 | 2 7 9 6 3 | 5 2 3 3 7 | 4 3 0 0 4 |
| 3 6 9 9 0 | 1 6 5 0 7 | 1 8 7 6 8 | 1 3 1 8 6 | 8 3 1 0 6 | 4 5 9 1 9 | 0 7 0 8 8 | 9 6 3 0 2 |
| 8 5 5 3 4 | 7 1 3 5 5 | 3 6 0 3 0 | 2 8 9 5 0 | 7 2 1 5 4 | 4 8 4 6 7 | 4 4 7 7 9 | 1 6 1 5 2 |
| 5 0 4 3 6 | 3 0 8 7 6 | 8 0 3 6 7 | 6 0 4 2 2 | 2 3 8 0 7 | 0 9 5 2 4 | 4 7 1 5 4 | 8 8 8 2 1 |
| 1 5 5 6 3 | 0 6 0 2 0 | 0 0 4 4 0 | 4 7 7 0 8 | 9 1 4 1 7 | 3 7 7 2 5 | 3 3 1 3 2 | 5 5 4 0 9 |
| 0 1 9 9 8 | 8 9 5 5 8 | 4 7 1 0 5 | 1 5 1 9 7 | 4 5 2 0 9 | 3 0 9 0 7 | 0 6 5 5 1 | 5 0 4 0 0 |
| 6 2 5 2 0 | 9 6 7 0 3 | 7 9 8 9 6 | 5 5 3 1 2 | 0 3 1 5 9 | 6 6 4 9 5 | 6 9 8 3 3 | 9 4 6 6 2 |
| 9 1 2 6 9 | 4 4 0 6 1 | 7 2 5 8 4 | 3 9 1 5 8 | 7 6 5 0 8 | 1 8 3 3 4 | 3 9 3 5 7 | 1 7 9 3 5 |
| 0 5 2 6 1 | 7 3 6 5 5 | 3 1 7 1 2 | 4 9 9 9 0 | 6 2 7 9 0 | 6 6 2 3 8 | 9 5 7 3 8 | 6 8 7 1 2 |
| 4 6 8 7 2 | 8 8 2 9 1 | 8 8 8 5 3 | 2 8 2 1 3 | 0 9 2 3 0 | 4 3 5 2 2 | 0 4 3 7 6 | 0 1 0 9 7 |
| 5 7 5 3 7 | 2 8 7 0 8 | 5 2 4 5 1 | 8 1 7 4 0 | 5 4 8 2 8 | 8 9 4 4 5 | 1 5 3 6 8 | 8 8 2 5 2 |
| 9 3 9 6 1 | 6 4 0 2 7 | 3 4 4 5 1 | 8 6 8 7 0 | 0 3 5 8 0 | 7 8 0 4 6 | 5 8 3 2 8 | 2 5 6 0 6 |
| 6 9 0 6 9 | 4 0 8 9 1 | 0 0 5 2 4 | 0 5 6 4 4 | 9 0 9 4 5 | 7 3 2 0 6 | 0 3 5 6 2 | 8 9 3 0 9 |
| 8 6 4 9 8 | 2 1 2 9 7 | 7 0 4 8 3 | 0 3 2 2 2 | 8 9 1 7 7 | 4 1 7 6 6 | 1 2 6 2 3 | 8 7 1 4 1 |
| 6 2 9 5 4 | 7 9 6 2 9 | 2 3 9 3 8 | 4 2 5 0 6 | 4 9 4 1 7 | 8 9 0 7 6 | 1 5 8 4 1 | 3 9 8 2 2 |
| 9 5 3 6 7 | 6 7 2 6 5 | 8 2 9 4 7 | 9 1 6 9 6 | 7 6 2 3 1 | 7 5 5 2 3 | 3 4 7 6 9 | 3 0 4 9 6 |
| 8 0 0 2 0 | 5 6 4 2 8 | 8 9 2 9 1 | 5 6 4 1 3 | 8 7 5 5 8 | 5 4 7 8 1 | 5 9 8 7 7 | 6 2 3 6 3 |
| 0 9 0 4 4 | 0 2 5 1 1 | 5 0 6 1 2 | 7 9 7 1 5 | 0 2 3 5 0 | 3 2 7 2 6 | 0 0 5 5 5 | 4 5 6 2 8 |
| 9 7 0 8 2 | 3 0 3 0 1 | 5 6 2 6 6 | 1 1 6 5 2 | 4 1 5 4 0 | 9 3 6 9 3 | 0 8 7 5 7 | 9 2 3 4 9 |
| 1 2 0 3 5 | 2 1 9 4 1 | 6 7 7 3 8 | 6 7 7 7 9 | 6 5 0 8 4 | 0 5 4 5 5 | 1 9 9 3 1 | 8 7 5 5 7 |
| 9 1 1 2 7 | 4 3 3 6 4 | 3 3 0 8 0 | 4 6 0 8 8 | 7 4 3 4 1 | 9 4 2 2 0 | 4 7 9 9 9 | 6 2 1 7 5 |
| 2 5 9 0 8 | 6 9 5 1 1 | 4 7 5 2 4 | 7 8 6 2 3 | 7 4 9 1 6 | 3 9 0 8 3 | 6 5 5 7 5 | 6 4 0 5 0 |
| 7 8 7 2 2 | 6 5 8 4 5 | 1 4 5 7 0 | 6 5 3 1 6 | 0 0 2 2 8 | 4 9 1 1 4 | 8 6 8 0 1 | 7 9 8 8 1 |
| 7 0 6 1 4 | 1 5 4 0 7 | 8 7 4 4 6 | 9 0 2 9 6 | 7 6 8 8 6 | 5 5 7 1 2 | 9 9 4 7 5 | 7 8 6 4 0 |
| 8 7 3 7 5 | 7 3 6 4 9 | 9 0 2 2 4 | 6 6 0 8 7 | 9 7 6 7 4 | 5 1 4 9 0 | 3 3 6 7 4 | 3 3 0 6 2 |
| 4 9 0 2 0 | 1 7 4 1 9 | 4 6 1 8 0 | 9 2 2 2 3 | 0 0 4 3 5 | 4 8 3 1 9 | 0 5 8 3 5 | 1 5 2 5 9 |
| 1 7 1 6 6 | 1 1 3 8 1 | 6 3 5 1 5 | 3 1 2 8 5 | 7 9 6 3 7 | 6 1 0 5 9 | 5 4 9 2 3 | 8 7 0 8 8 |
| 7 6 8 0 0 | 5 7 1 8 7 | 9 3 2 4 8 | 6 3 7 9 2 | 1 6 9 9 4 | 0 7 9 7 2 | 7 5 3 3 7 | 2 7 5 5 9 |
| 6 4 9 0 8 | 2 2 9 7 1 | 7 0 8 3 7 | 3 0 0 4 0 | 5 5 4 9 7 | 6 8 4 9 6 | 3 1 0 1 7 | 1 0 3 2 9 |
| 9 6 5 4 4 | 6 0 4 5 3 | 2 1 9 5 6 | 5 6 3 8 5 | 5 3 3 6 6 | 3 5 7 1 7 | 4 3 4 0 2 | 8 0 3 9 7 |
| 2 0 7 4 6 | 0 2 6 5 6 | 8 2 8 9 0 | 9 5 2 5 6 | 7 0 3 4 6 | 7 8 1 1 8 | 5 3 8 8 5 | 6 6 5 9 4 |
| 0 9 8 9 2 | 0 3 8 5 8 | 7 9 4 6 6 | 2 5 2 3 0 | 5 1 6 9 6 | 2 5 1 9 1 | 5 1 1 9 2 | 7 5 9 7 8 |
| 8 1 8 5 8 | 6 7 6 5 3 | 8 5 1 6 7 | 0 3 5 5 5 | 9 9 3 2 7 | 0 2 3 5 4 | 2 0 9 4 1 | 6 7 5 0 9 |
| 8 4 7 2 0 | 3 2 5 2 3 | 1 7 9 7 9 | 4 3 3 3 4 | 1 6 6 9 5 | 0 8 9 2 1 | 9 7 8 4 9 | 8 2 6 5 2 |
| 7 2 5 8 7 | 4 0 1 3 5 | 6 1 0 6 7 | 0 3 4 8 6 | 1 4 8 2 3 | 2 9 2 0 0 | 9 2 7 1 9 | 0 4 4 7 4 |
| 5 6 1 8 4 | 3 4 9 8 3 | 7 6 3 4 1 | 1 6 5 5 8 | 5 4 5 0 1 | 9 0 5 7 9 | 3 6 4 2 3 | 2 1 8 8 5 |
| 2 7 4 5 0 | 9 0 6 8 0 | 5 8 3 8 8 | 7 6 5 8 6 | 1 1 4 6 3 | 3 5 7 2 9 | 0 8 4 9 3 | 7 4 7 9 4 |
| 2 2 1 4 5 | 8 7 1 0 3 | 2 7 7 3 1 | 2 7 2 6 5 | 5 8 0 3 9 | 8 6 7 8 7 | 3 9 2 7 2 | 5 9 7 1 7 |
| 4 0 8 9 5 | 9 2 1 6 0 | 8 0 2 5 2 | 2 4 6 2 0 | 3 2 4 7 1 | 5 4 7 6 3 | 4 5 1 1 5 | 6 7 2 0 2 |
| 6 8 7 4 8 | 5 8 2 2 8 | 6 1 0 6 1 | 8 4 6 3 1 | 7 4 5 2 8 | 6 5 7 3 4 | 4 4 9 5 9 | 4 0 2 1 2 |
| 1 9 1 1 3 | 6 2 3 9 8 | 1 7 0 3 3 | 6 2 0 3 9 | 6 4 8 1 0 | 3 6 1 4 9 | 7 2 6 3 2 | 7 1 2 9 4 |
| 8 2 7 3 7 | 6 6 6 8 8 | 7 6 9 2 3 | 5 6 4 1 6 | 7 6 1 1 6 | 6 0 8 7 0 | 9 1 0 6 4 | 2 6 8 6 7 |
| 2 9 1 5 1 | 0 7 6 8 1 | 8 7 9 4 7 | 0 5 5 1 5 | 1 5 9 9 8 | 8 5 1 7 4 | 0 9 1 0 1 | 6 4 3 3 1 |
| 6 6 2 2 2 | 9 3 9 6 5 | 4 1 5 8 6 | 4 8 4 1 7 | 4 1 7 7 2 | 2 5 7 8 4 | 0 5 0 9 4 | 7 6 4 1 9 |
| 1 0 2 3 4 | 5 9 0 2 0 | 7 0 8 7 2 | 7 4 0 9 9 | 4 5 7 8 8 | 8 3 1 3 7 | 6 3 2 2 2 | 1 2 5 2 2 |

15. Suppose a population has a distribution with $\mu = 72$ and $\sigma = 8$.

- (a) If we know nothing about the original population distribution and random samples of size $n = 16$ are selected, why can't we say anything about the \bar{x} distribution of sample means?
- The sample size is too small to be certain what the resulting \bar{x} distribution will look like. In order to apply the Central Limit Theorem, which asserts that the \bar{x} distribution is approximately normal, the sample size needs to be $n \geq 30$.
- (b) If we know that the original population distribution is normal, then what can we say about the \bar{x} distribution of random samples of $n = 16$. In this case, find $P(68 \leq \bar{x} \leq 73)$.
- When the original population distribution is normal, then so will the \bar{x} distribution, regardless of the size of the samples.
- The \bar{x} distribution still has a mean $\mu_{\bar{x}} = \mu = 72$, but the standard deviation is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{16}} = 2$.
- $P(68 \leq \bar{x} \leq 73) = \text{normalcdf}(68, 73, 72, 2) \approx 0.6687124058$.
16. The heights of 18-year-old men are approximately normally distributed, with a mean of 68 inches and a standard deviation of 3 inches.
- (a) What is the probability that an 18-year-old man selected at random is between 67 and 69 inches tall?
- $P(67 < x < 69) = \text{normalcdf}(67, 69, 68, 3) \approx 0.2611171926$
- (b) If a random sample of nine 18-year-old men are selected at random, what is the probability that the mean height of the sample \bar{x} is between 67 and 69 inches?
- In this case, it is not the value of the random variable x that we are considering, but rather the value of the mean \bar{x} of samples of size $n = 9$.
- Because the distribution of x is normal, the distribution of \bar{x} will be too. Further, the mean of the \bar{x} distribution is still 68, but the standard deviation of that distribution is $\frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{9}} = 1$.
- $P(67 < \bar{x} < 69) = \text{normalcdf}(67, 69, 68, 1) \approx 0.6826894809$
- (c) Explain why the probability in part (b) is so much higher than the probability in part (a) even though they both refer to the same interval of heights (67 to 69 inches).
- The standard deviation of the sample mean distribution is smaller than that of the original distribution, so there is less variance in the values. That is, the mean height of a randomly selected sample of men is more likely to be close to the mean height of all men than a single randomly selected man's height is.
17. Let x be a random variable that represents white blood cell count per cubic milliliter of whole blood. Assume that x has a distribution that is approximately normal with a mean of $\mu = 7500$ and estimated standard deviation of $\sigma = 1750$. A test result of $x < 3500$ is an indication of leukopenia. This indicates bone marrow repression that may be the result of a viral infection.
- (a) What is the probability that, on a single test, x is less than 3500?
- $P(x < 3500) = \text{normalcdf}(-1\text{E}99, 3500, 7500, 1750) \approx 0.01113549 \approx 1.11\%$
- (b) Suppose a doctor uses the average \bar{x} for two tests taken about a week apart. What can we say about the probability distribution of \bar{x} ? What is the probability that $\bar{x} < 3500$?
- Because x has a distribution that is approximately normal, the distribution of the averages of the samples (even with size as small as $n = 2$) is also normal. The mean of the distribution of averages is still 7500, however, the standard deviation is reduced to $\frac{\sigma}{\sqrt{n}} = \frac{1750}{\sqrt{2}}$.
- $P(\bar{x} < 3500) = \text{normalcdf}(-1\text{E}99, 3500, 7500, 1750/\sqrt{2}) \approx 0.0006135861 \approx 0.061\%$

- (c) Repeat part (b) but with $n = 3$ tests.

Again, the distribution of the averages of three test results is normal, with mean 7500 and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{1750}{\sqrt{3}}$.

$$P(\bar{x} < 3500) = \text{normalcdf}(-1\text{E}99, 3500, 7500, 1750\sqrt{3}) = 0.00003763633 \approx 0.0038\%$$

- (d) How did the probabilities change as n increased? What do such results imply about a patient that has $\bar{x} < 3500$ for 3 those tests?

Although it is possible for an isolated random test to indicate a low count when absolutely nothing is wrong, the greater n is, it becomes less and less likely that the average of the tests will be less than 3500 when absolutely nothing is wrong.

According to part (c), the likelihood that the average of 3 tests is less than 3500 just by random chance is only 0.00003763633. So, for a patient with such results, the probability is extremely high that they do, in fact, have leukemia.

18. Assume that IQ scores are normally distributed, with a standard deviation of 15 points and a mean of 100 points. If 10 people are chosen at random, what is the probability that the sample mean of their IQ scores will not differ from the population mean by more than 2 points?

In this case, we are looking for the probability that the sample mean \bar{x} falls between 98 and 102 points. Because the distribution of IQ scores is normal, the Central Limit Theorem can be applied with $\mu_{\bar{x}} = 100$ and $\sigma_{\bar{x}} = \frac{15}{\sqrt{10}}$.

$$\text{So, } P(98 \leq \bar{x} \leq 102) = \text{normalcdf}(98, 102, 100, \frac{15}{\sqrt{10}}) \approx 0.3267099507 \approx 32.67\%.$$

19. A large tank of fish from a hatchery is being delivered to a lake. The hatchery claims that the mean length of fish in the tank is 15 inches, and the standard deviation is 2 inches. A random sample of 36 fish is taken from the tank. Let \bar{x} be the mean length of the sample. What is the probability that \bar{x} is within 0.5 inches of the claimed population mean?

Although there is no indication that the distribution of lengths of the fish is normal, the sample size is greater than 30, so the Central Limit Theorem applies. So, we can apply the **normalcdf** function with $\mu_{\bar{x}} = 15$ and $\sigma_{\bar{x}} = 2/\sqrt{36} = 1/3$.

$$\text{normalcdf}(14.5, 15.5, 15, 1/3) \approx 0.8663855426 \approx 86.64\%$$