To Null or Not to Null

Imagine that you are a member of a jury in a criminal trial. What kinds of mistakes are possible in your verdict? In this module, you explore how statistics can help you analyze this situation.

Cliff Bara • Pat Mauch • Lisa Wood

To Null or Not to Null

Introduction

Most of the choices you make each day contain some degree of uncertainty. When things go wrong, the consequences of a mistake can range from almost insignificant to very grave. The process of using statistics to help make these choices—and measure the consequences—is known as **hypothesis testing**. Although the use of statistics cannot guarantee that you will never make a mistake, it can allow you to measure the risk of an error.

Discussion

a. Consider each of the following scenarios. For each one, describe the possible consequences of a wrong decision.

1. You are a member of a jury. The defendant in the case has been accused of murder. You must decide if this person is innocent or guilty.

2. You are the president of a tire manufacturing company. A consultant has recommended that you increase the tread life of your tires. You must decide whether to approve the additional spending required for this upgrade.

3. You are editor of a high school newspaper. On the recent Scholastic Aptitude Test (SAT), the senior class scored slightly higher than the national average on the mathematics portion. A reporter has submitted an article claiming that the school's seniors are better at mathematics than students nationally. You must decide whether or not to publish the article.

b. Which of the scenarios described above could be analyzed statistically? If an analysis is possible, explain briefly how it might be done.

c. When deciding how to treat a patient's illness, doctors often rely on test results. Many of these tests are not 100% accurate. Given this fact, why do doctors use such tools?

d. During Olympic and other world-class competitions, athletes must undergo testing for drugs and other banned substances.

1. If an athlete tests positive for a specific substance, does this guarantee that he or she has used the drug?

2. How would you expect a rules committee to react to the news of a positive drug test?

Mathematics Note

Statisticians often make hypotheses or claims about the parameters of a population, then use sampling techniques to test their claims. If a researcher assumes that a population parameter has a specific value, then a hypothesis can be formed about the consequences of that assumption.

In statistical analysis, there are two types of hypotheses. A **null hypothesis** (**H_0**) is a statement about one or more parameters. The **alternative hypothesis** (**H_a**) is the statement that must be true if the null hypothesis is false. The null hypothesis usually involves a claim of no difference or no relationship. In many situations, the null hypothesis and alternative hypothesis are negations of each other, but this is not necessarily the case.

For example, consider a study of tread life for two types of tires: A and B. In this case, the null hypothesis would be that there is no difference in tread life between tire A and tire B, or H_0: $A = B$. An alternative hypothesis could be that the two tires have different tread lives, or H_a: $A \neq B$. If the study revealed that the tread life on tire A appear to be longer than tire B, the alternative hypothesis could be H_a: $A > B$.

e. Suppose that a consumer group wanted to test a manufacturer's claim that its light bulbs have a mean life of 1000 hr. The study team formulates the null hypothesis "H_0: $\mu = 1000$," where μ represents the population mean.

1. What is the negation of this null hypothesis?
2. In this situation, the study team decides to use the alternative hypothesis "H_a: $\mu < 1000$." Why do you think these researchers did not use the negation of the null hypothesis as their alternative hypothesis?

f.

1. Suggest null and alternative hypotheses for a situation in which an athlete undergoes a drug test.
2. If your null hypothesis (H_0) is false, what type of evidence would you expect to observe in the drug test?

Activity 1

Statisticians are seldom 100% confident of their findings. In this activity, you explore how uncertainty affects hypothesis testing.

Mathematics Note

A **hypothesis test** may consist of the following steps.

- State null and alternative hypotheses about a parameter of a population.
- If the null hypothesis is true, predict what this implies about a sample of the population.
- Take a sample of the population and compare the results with your prediction.
- If the results are inconsistent with the prediction, then you can conclude, with some level of certainty, that the null hypothesis is false and, therefore, reject it.
- If the results are consistent with the prediction, you fail to reject the null hypothesis. The failure to reject the null hypothesis does not guarantee that the null hypothesis is true, only suggests that it might be true.

Exploration

In the following exploration, you use a population of coins to investigate how uncertainty affects your interpretation of test results.

a. Place a penny in each of 18 envelopes. Place a nickel in one additional envelope, and a quarter in another envelope. **Note:** For the remainder of this exploration, these will be referred to as the "test envelopes."

b. Randomly select one of the test envelopes. Then obtain an envelope from your teacher containing an unknown coin. In this situation, the null and alternative hypotheses can be stated as follows:

H_0: The unknown coin is a penny .

H_a : The unknown coin is not a penny .

1. If the unknown coin is a penny, what would you expect to occur when this envelope and a test envelope are placed on opposite sides of a balance? What percent of the time would you expect the test envelope to contain a penny?
2. What would you expect to occur if the unknown coin is not a penny? What percent of the time would you expect the envelope to contain the nickel or quarter?

c. Place the test envelope on one side of a balance. Place the envelope containing the unknown coin on the opposite side of a balance. Use your observations to decide whether to reject, or fail to reject, the null hypothesis.

d. Describe the conclusions you can make as a result of your hypothesis test.

Discussion

a. Do you think that the unknown coin is a penny? Explain your response.

b. What is the probability of selecting a test envelope that contains a penny?

c. **1.** If the two envelopes in Part **c** of the exploration balance, can you be sure that the coin is a penny?

2. If the two envelopes do not balance, can you be sure that the coin is not a penny?

d. Given only the results of a single test, is there any way to remove the uncertainty from your conclusions? Explain your response.

Mathematics Note

Whenever a sample is taken from a population about which there is some uncertainty, it is possible that the sample is not representative of the population. Therefore, when performing a hypothesis test by sampling, it is always possible to make an incorrect decision about the null hypothesis H_0. These two possible errors, rejecting a true null hypothesis and failing to reject a false null hypothesis, are shown in Figure **1** below.

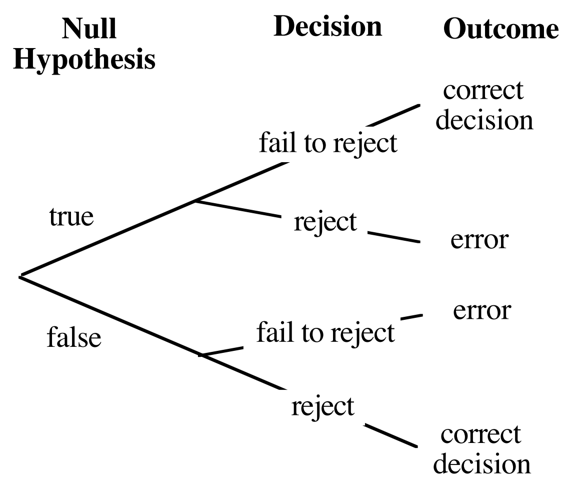


Figure 1: Possible errors for hypothesis test

In the exploration, for example, it is possible to draw a test envelope that contains a nickel rather than a penny. If the null hypothesis states that the unknown coin is a penny, then you would expect the two envelopes to balance. If the test envelope contains a nickel, however, it will not balance with an envelope containing a penny. This would lead you to reject a true null hypothesis.

e. **1.** Why does the rejection of the null hypothesis result in the acceptance of the alternative hypothesis?

2. Does acceptance of the alternative hypothesis guarantee that it is true?

f. **1.** If you fail to reject a null hypothesis, does that prove that it is true? Use the tree diagram in Figure **1** to support your answer.

2. If you reject a null hypothesis, does that prove that it is false? Explain your response.

Assignment

1.1 While studying a developing economy, researchers formulated the following null hypothesis: "The mean annual salary in the population is at least $10,000." What is the alternative hypothesis in this situation?

1.2 A mail-order catalog claims that customer satisfaction is guaranteed. Write the null and alternative hypotheses that you would use in testing this claim.

1.3 Explain why it is not typically possible to guarantee a correct decision when failing to reject a null hypothesis.

1.4 Consider a set of 20 test envelopes: 19 contain a penny and 1 contains a quarter. Using the procedure described in the exploration and an envelope containing an unknown coin, you test the null hypothesis: "The unknown coin is a penny." If the two envelopes balance, how certain can you be of your conclusion?

* * * * *

1.5 In 1994, a U.S. state began collecting a tax on tourism. The state's current governor wants to compare spending by tourists before and after the tax was enacted.

a. Write null and alternative hypotheses for this situation.

b. Describe the types of errors that might occur when testing the null hypothesis in Part **a**.

1.6 A standard deck of 52 playing cards contains 26 red cards and 26 black cards. Suppose that you wanted to determine the proportion of red cards in an unknown deck by sampling. In this situation, the null hypothesis might be: "The proportion of red cards in the deck is 0.5."

a. What is the alternative hypothesis that must be accepted if the null hypothesis is rejected?

b. To test your null hypothesis, you select a random sample of 6 cards from the deck. What outcomes might lead you to suspect that the proportion of red cards in the deck is not 0.5?

c. What is the probability of drawing a random sample of 6 cards, all of which are the same color, from a standard deck?

d. If your sample of 6 cards from this deck were all the same color, would you reject the null hypothesis? Explain your response.

* * * * * * * * * *

Activity 2

The counselor at John F. Kennedy High School has just reported the results of the SAT to 32 seniors. Their scores on the mathematics portion are normally distributed, with a mean of 529. One senior, Dena, scored 610. Although Dena knows she performed better than the mean, she would like to know approximately how many of her classmates she outscored.

In Activity **1**, you formed decisions about hypotheses with various levels of certainty. In this activity, you use a normal curve to determine how an individual observation, such as a test score, compares to a population parameter, such as the mean. You also discover how to compare individual observations from different populations, such as scores from two different tests.

Mathematics Note

The curve that describes the shape of a normal distribution is the **normal curve**. The equation of the curve that models a normally distributed population depends on the mean and standard deviation of the population. The general equation for a normal curve, where μ represents the population mean and σ represents the population standard deviation, is:

$$y = \frac{1}{\sigma\sqrt{2\pi}}e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2}$$

Recall that both π and e are constants with values of approximately 3.14159 and 2.71828, respectively.

For example, Figure **2** shows the normal curve for a population of test scores where $\mu = 80$ and $\sigma = 5$:

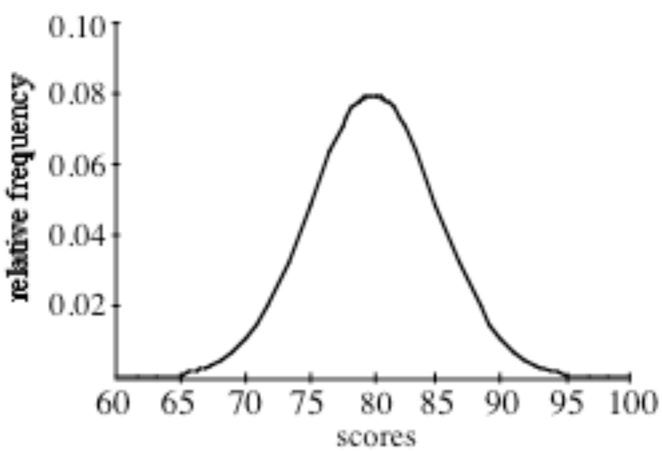


Figure 2: A normal curve

A normal curve describes a continuous probability distribution. As in all such distributions, the area between the horizontal axis and the curve is 1.

Exploration

a. The scores on a recent English examination at JFK High School are normally distributed, with a mean of 80 ($\mu = 80$) and a standard deviation of 5 ($\sigma = 5$). Recall that the **standard deviation** (σ) is a measure of the variability, or spread, within a population. Create a graph that models the distribution of these scores by substituting these values into the general equation for a normal curve.

b.
1. Locate the point on the x-axis that corresponds with the mean. Draw a vertical segment from the x-axis to the curve at this point.

2. Determine the percentage of the area between the x-axis and the curve that lies to the left of the segment in Step **1**.

c.
1. Draw vertical segments from the x-axis to the curve at $x = \mu - 1\sigma$ and $x = \mu + 1\sigma$.

2. Find the percentage of the area between the x-axis and the curve that lies between these two segments.

d.
1. Draw vertical segments from the x-axis to the curve at $x = \mu - 2\sigma$ and $x = \mu + 2\sigma$.

2. Find the percentage of the area between the x-axis and the curve that lies between these two segments.

e. Repeat Parts **a–d** using two different values for σ, when $\mu = 80$. Record your observations.

f. Repeat Parts **a–d** using two different values for μ, when $\sigma = 5$. Record your observations.

g. The scores on a recent physics exam at JFK High School are normally distributed with $\mu = 72$ and $\sigma = 3$. Repeat Parts **a–d** for this set of scores.

Discussion

a. When using a continuous curve like the one in Figure **2** to approximate the distribution of test scores, what assumptions are made about the scores?

b. How does the value of σ for a particular normal distribution affect the shape of the corresponding normal curve?

c. How does the value of σ for a particular normal distribution affect the percentage of area above the x-axis and below the normal curve in each of the following intervals?

1. $[\mu - 1\sigma, \mu + 1\sigma]$

2. $[\mu - 2\sigma, \mu + 2\sigma]$

d. How does the value of μ for a particular normal distribution affect the shape of the corresponding normal curve?

e. Describe how the percentages you determined in Parts **b–d** of the exploration relate to probabilities.

Mathematics Note

The predictable variability of individual observations is summarized by the **68–95–99.7 rule**, which states that approximately 68% of the area between the *x*-axis and the normal curve is within 1 standard deviation of the mean, approximately 95% is within 2 standard deviations of the mean, and approximately 99.7% is within 3 standard deviations of the mean. This rule is represented graphically in Figure **3**.

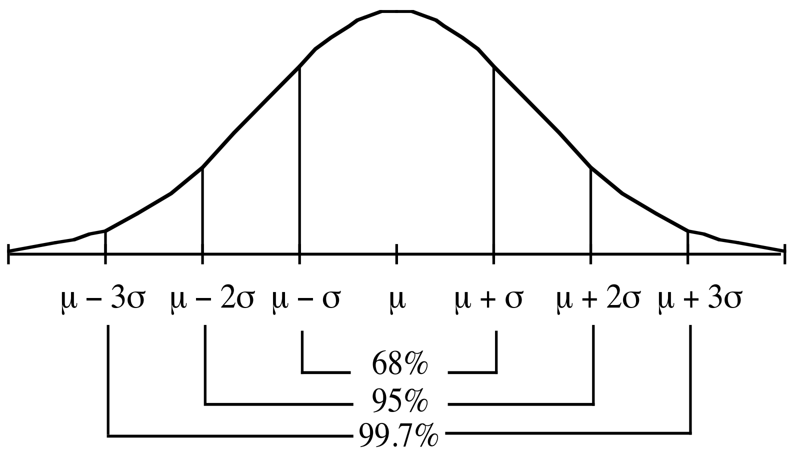


Figure 3: The 68-95-99.7 rule

The 68–95–99.7 rule can be used to make predictions about populations that are normally distributed. For example, consider a set of test scores that are normally distributed with a mean of 80 and a standard deviation of 5. About 68% of these scores would fall between $80 - 5 = 75$ and $80 + 5 = 85$. Similarly, about 95% of the scores would fall in the interval [70, 90], while about 99.7% would fall in the interval [65, 95].

f. If a set of scores is normally distributed, describe how the 68-95-99.7 rule can be used to estimate the probability that a random score from the set is:

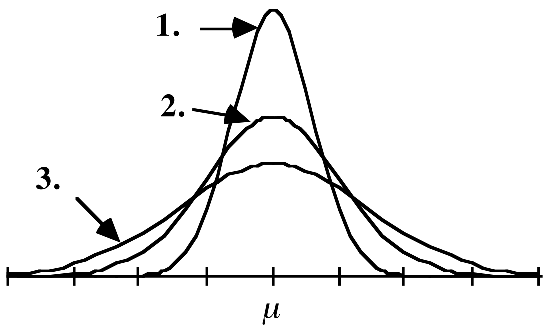
1. contained in the interval $[\mu, \ \mu + 1\sigma]$

2. contained in the interval $[\mu - 2\sigma, \ \mu]$

3. is not contained in the interval $[\mu - 2\sigma, \ \mu + 2\sigma]$

g. Why does the area between the *x*-axis and the curve for all continuous probability distributions equal 1?

h. The graph below shows three normal curves with the same mean (μ), but different standard deviations.



Which curve represents the population with the largest standard deviation? Justify your response.

i. Consider a set of scores with a mean of μ and a standard deviation of σ. Describe how to determine the number of standard deviations a score of x is above or below the mean.

Mathematics Note

Any value x from a normally distributed population with mean μ and standard deviation σ can be represented by a ***z*-score**. A z-score describes the number of standard deviations that the value is above or below the mean. The formula for determining a z-score is shown below:

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

Because the percentage of values lower than x in a normally distributed population can provide useful information, these percentages are commonly available in books and tables. A portion of such a table is shown in Table **1**.

Table 1: Portion of a *z*-score table

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633

In this table, the left-hand column lists z-scores from 1.5 to 1.7 by tenths. The hundredths place for each of these values is displayed in the top row. Each four-digit decimal value in the table represents the area between the x-axis and a normal curve that lies to the left of the corresponding value of x. **Note:** A complete table of z-scores and their corresponding areas appears at the end of this module.

For example, consider a set of normally distributed test scores with a mean of 80 and a standard deviation of 3. In this situation, a test score of 85 can be represented by the following z-score: $z = (85 - 80)/3 \approx 1.67$. From the table, a z-score of 1.67 corresponds with a decimal of 0.9525. This means that approximately 95.25% of the test scores in this population are below a test score of 85. In other words, the probability that a test score from this population is less than 85 is approximately 95.25%. Conversely, the probability that a test score from this population is greater than 85 is $1 - 0.9525 = 0.0475$, or approximately 4.75%.

j. Is it possible for z to be negative? If so, when? If not, why not?

k. Dena scored 610 on the mathematics portion of the SAT. The average score for the 32 seniors was 529, with a standard deviation of 125. About how many seniors did Dena outscore?

l. Consider two sets of test scores. In set A, the mean is 70 and the standard deviation is 5. In set B, the mean is 72 and the standard deviation is 3.

1. Based on the number of standard deviations above the mean, which is the better test score: a 76 in set A or a 76 in set B?
2. Based on the percentage of scores below 76, which is the better score?

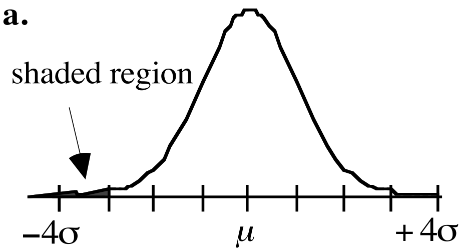
m. A set of 50 English examination scores is normally distributed with a mean of μ and a standard deviation of σ. Describe how to find the probability that a score selected at random from this set is less than or equal to 64, given each of the following values for μ and σ.

1. $\mu = 71$ and $\sigma = 4$
2. $\mu = 75$ and $\sigma = 5$
3. $\mu = 48$ and $\sigma = 5$

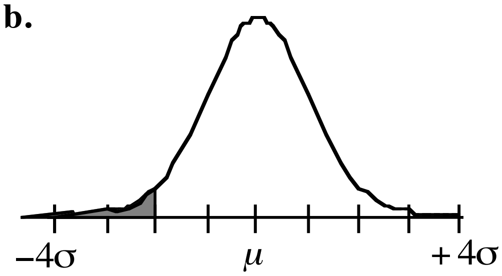
Assignment

2.1 Determine the percentage of the total area under the curve represented by the shaded region(s) in each of the following graphs.

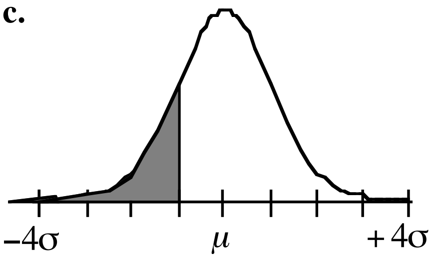
a.



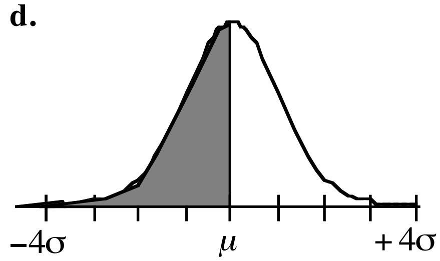
b.



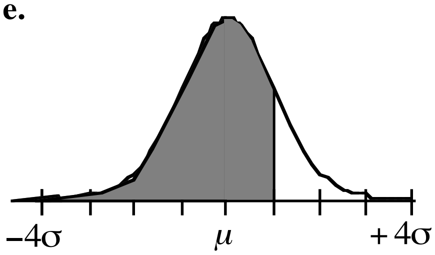
c.



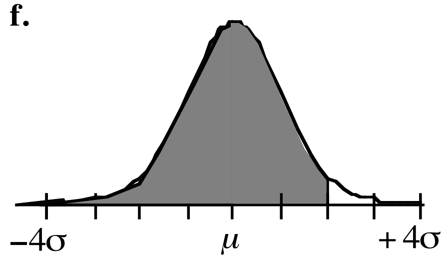
d.



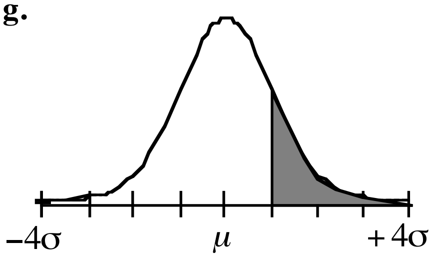
e.



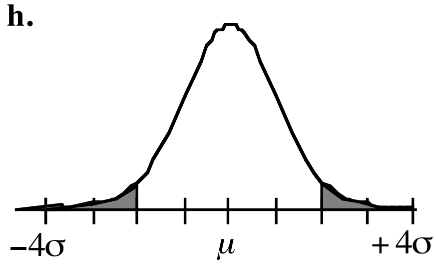
f.



g.



h.



2.2 Determine the area between the x-axis and a normal curve for $x < \mu - 1.25\sigma$.

2.3 Use the tables for Area under Normal Curve found at the end of this module to determine eachof the following:

a. Find the area between the x-axis and a normal curve that lies to the left of each of the following values:

1. $\mu - 3\sigma$

2. $\mu - \sigma$

3. $\mu + 2\sigma$

b. Find the area between the x-axis and a normal curve that lies to the right of each of the following values:

1. $\mu - 2\sigma$

2. μ

3. $\mu + \sigma$

4. $\mu + 3\sigma$

2.4 Rolf earned an 82 on an English exam that was normally distributed with $\mu = 80$ and $\sigma = 5$. Dena received a 77 on a physics exam that was normally distributed with $\mu = 72$ and $\sigma = 3$.

Rolf claims, “My score is better because 82 is greater than 77.” Dena says, “No way! My score is better because the physics exam was tougher.” Defend either Rolf’s or Dena’s position.

2.5 Consider a set of normally distributed test scores with $\mu = 79$ and $\sigma = 5$. What is the probability that a randomly selected score from this set is:

a. less than 70?

b. less than 85?

c. between 74 and 84?

d. between 75 and 90?

e. greater than 83?

* * * * *

2.6 In 1993, 1,044,465 students took the mathematics portion of the SAT. The test scores were normally distributed with $\mu = 478$ and $\sigma = 125$.

a. What percentage of students scored below 500 on the exam? (This is the percentile rank associated with a score of 500.)

b. Approximately how many people scored above 500 on the exam?

c. To be considered for a distinguished scholarship at the local university, a candidate must score in the 90th percentile or better (without rounding) on the mathematics portion of the SAT. What score must a candidate obtain to be considered for this award?

2.7 Is it possible for two sets of examination scores to exist in which a score of 90 on one test is not as good as a score of 47 on the other? Justify your response with examples.

2.8 A set of test scores is normally distributed with $\mu = 72$ and $\sigma = 3$.

a. If one test score is selected at random from this set, what is the probability that this score is:

1. greater than 74?
2. greater than 82?
3. greater than 68?
4. greater than 70?

b. Maeve scored 74 on this test. She claims she performed significantly better than the others who took the exam. Support or refute her claim.

2.9 The mean body temperature of a healthy person is 37˚C. Assume that body temperatures are normally distributed about this mean, with a standard deviation of 0.23˚C. Within what range of values would you expect the temperature of 95% of healthy people fall?

* * * * * * * * * *

Activity 3

In Activity **2**, Dena used a normal curve to compare her individual test score with the mean of 32 scores from her class. However, this set of scores did not include the scores of all the seniors at John F. Kennedy High School.

In an article for the school newspaper, Dena would like to claim that this year's seniors did significantly better on the SAT than last year's seniors. Can she use the mean of 32 scores to estimate the mean for the entire class?

Mathematics Note

The **sampling distribution of sample means** contains the means ($\bar{x}$) of *all* possible samples of size *n* from a population.

The mean of the sampling distribution of sample means, denoted by $\mu_{\bar{x}}$, equals the population mean μ.

The standard deviation of the sampling distribution, denoted by $\sigma_{\bar{x}}$, equals $\sigma/\sqrt{n}$, where σ is the population standard deviation and *n* is the sample size. When σ is unknown, the standard deviation of the sample (*s*) may be used as an estimate of σ.

The **central limit theorem** states that, even if the population from which samples are taken is not normally distributed, the sample means tend to be normally distributed. The approximation to the normal curve becomes more accurate as the sample size n increases. For $n \geq 30$, the distribution of sample means can be modeled reasonably well by a normal curve. This requirement is not necessary if the population from which samples are taken is normally distributed.

For example, suppose you take a sample of 32 boxes of cereal and determine the mean of their masses. Since $n \geq 30$, the central limit theorem applies and the sampling distribution of sample means can be modeled by a normal curve. Using the 68–95–99.7 rule, approximately 68% of the sample means fall within 1 standard deviation of the population mean μ; approximately 95% of the sample means fall within 2 standard deviations of μ; and approximately 99.7% of the sample means fall within 3 standard deviations of μ.

Discussion 1

a. How do $\bar{x}$, μ, and $\mu_{\bar{x}}$ compare?

b. How do σ and $\sigma_{\bar{x}}$ compare?

c.
1. Consider a random sample of 50 test scores taken from a population of scores. Should the 68–95–99.7 rule be used to compare the sample mean to the population mean? Why or why not?
2. Consider a random sample of 20 test scores taken from a population of scores. Should the 68–95–99.7 rule be used to compare this sample mean to the population mean? Why or why not?

d. Consider a sampling distribution with a mean of $\mu_{\bar{x}}$ and a standard deviation of $\sigma_{\bar{x}}$. Describe how to determine the probability that a sample mean $\bar{x}$ selected at random from this distribution is:

1. more than 1.6 standard deviations below the mean
2. more than 2.3 standard deviations above the mean.

Exploration

The scores for 32 of this year's seniors had a mean of 529 and a standard deviation of 125. Assuming that these students represent a random sample of the class, Dena can use their mean and standard deviation to estimate the mean and standard deviation of the scores for the entire class.

The mean score on the mathematics portion of the SAT for last year's seniors was 478. Dena would like to claim that this year's seniors did significantly better than last year's class.

Through hypothesis testing, you can determine if the difference between a sample mean and a population mean is expected or not. If the difference is greater than what might be expected due to the predictable variability among samples, the difference is said to be "significant."

a. Formulate the null and alternative hypotheses for the situation described above.

b. Just as any value x from a normally distributed population with mean μ and standard deviation σ can be represented by a z-score, so can any value $\bar{x}$ from a normally distributed population with mean $\mu_{\bar{x}}$ and standard deviation $\sigma_{\bar{x}}$. In this case, the z-score (denoted by $z_{\bar{x}}$) can be determined as follows:

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Calculate the z-score for the sample mean of 529. Use the sample standard deviation of 125 as an estimate of the population standard deviation.

Mathematics Note

The **significance level** of a hypothesis test is an arbitrarily assigned probability that distinguishes a significant difference from a chance variation. Traditionally, researchers use significance levels of 0.10, 0.05, or 0.01.

Using a 0.10 significance level, for example, requires that the sample give evidence against the null hypothesis so strong that this evidence would occur no more than 10% of the time, assuming the null hypothesis is true. In other words, the chances of making the error of rejecting a true null hypothesis are less than 0.10, or 10%.

The **critical region** represents the set of all values that would lead a researcher to reject the null hypothesis. For example, suppose that a researcher has obtained a positive z-score for a statistic and selected a 0.10 significance level. From a table, the positive z-score associated with an area of 0.10 is 1.28. This value defines the boundary of the critical region, as shown in Figure **4**.

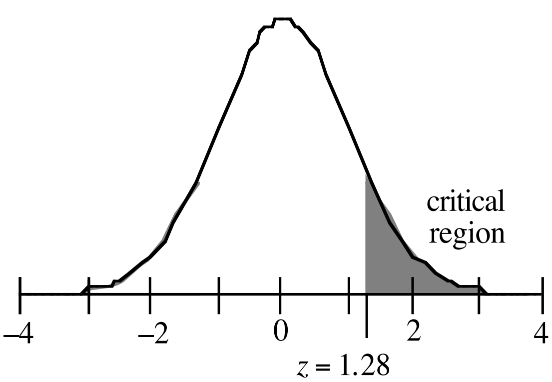


Figure 4: Critical region for 0.10 significance level

If the z-score of the individual observation (or sample mean) falls in the critical region, the researcher should reject the null hypothesis. This indicates that the difference between the statistic and the population parameter is not due to predictable variability at the given level of significance.

For example, suppose that a sample of 50 test scores has a mean of 81.5. Is this sample significantly better than a population with a mean of 80 and a standard deviation of 5? In this situation, the null hypothesis is that there is no difference between the mean of the population from which the sample is taken and 80.

Since the sample size is greater than 30, the central limit theorem applies, and the z-score can be calculated as follows:

$$z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{81.5 - 80}{5 / \sqrt{50}} \approx 2.11$$

The critical region for a 0.10 significance level is $z > 1.28$. This indicates that z-scores greater than 1.28 would occur by chance less than 10% of the time. Since the z-score of 2.11 is greater than 1.28, it falls in this critical region. The researcher should reject the null hypothesis at the 0.10 significance level.

If the researcher wanted to test at the 0.01 significance level, the critical region is $z > 2.33$. This indicates that z-scores greater than 2.33 would occur by chance less than 1% of the time. Since 2.11 is less than 2.33, the z-score falls outside the critical region. The researcher should fail to reject the null hypothesis at a 0.01 significance level.

c. Using a 0.05 significance level, determine the critical region for Dena's hypothesis test.

d. Determine whether Dena should reject, or fail to reject, the null hypothesis. Justify your choice.

e. Based on the results of the hypothesis test, what conclusion can you make?

Discussion 2

a. How does the formula for the z-score of an individual observation (x) compare to the formula for the z-score of a sample mean ($\bar{x}$)?

b. How did you determine the null and alternative hypotheses in Part **a** of the exploration?

c. Could Dena use the null hypothesis in the exploration to show that this year's class did the same as last year's class? Justify your response.

d. Describe how you determined the critical region for a hypothesis test at a 0.05 significance level.

e. In some hypothesis tests, the critical region is determined by two values, not one. This occurs when the null hypothesis states that a parameter equals a given value.

For example, to test the null hypothesis, H_0: $\mu = 5$, a researcher must consider possible values for μ that are both above and below 5. If a 0.05 significance level is desired, then the critical region is defined by $z = 2$ and $z = -2$, since approximately 95% of possible sample means fall within 2 standard deviations of μ.

1. Describe a graph of the critical region in this situation.
2. If the z-score for the statistic falls in the critical region, what should the researcher decide?

f. In what kinds of situations would a researcher select a 0.01 significance level rather than a 0.05 or 0.10 significance level?

g. How did you justify your decision to reject, or fail to reject, the null hypothesis?

h. Do you think that Dena should claim that this year's seniors did significantly better than last year's class on the SAT? Explain your response.

Assignment

3.1 In 1993, 2234 students in the state of Montana took the mathematics portion of the SAT. A random sample of 100 of these students had a mean score of 516 with a standard deviation of 125.

The mean score on this test for students around the nation was 478. Can the governor of Montana claim, with 90% certainty, that Montana students who took this test scored significantly better than students nationally? To make this decision, complete the following steps.

a. State the null and alternative hypotheses in this situation.

b. Determine $z_{\bar{x}}$ for the sample statistic.

c. Decide whether to reject, or fail to reject, the null hypothesis. Justify your reasoning.

d. Use the results of the hypothesis test to state a conclusion.

3.2 The mean score on the mathematics portion of the SAT for one class of 32 students was 529, with a standard deviation of 110. The mean score statewide was 516. A parent would like to know, with 95% certainty, if this class did significantly better on the test than the rest of the students in the state.

a. Formulate null and alternative hypotheses for this situation.

b. Find the z-score for the sample mean.

c. Decide whether to reject, or fail to reject, the null hypothesis. Justify your reasoning.

d. Explain what your decision means in this situation.

3.3 The following table shows the scores for 40 seniors at Washington High School on the mathematics portion of the SAT.

530	610	520	440	490	530	500	480
770	530	520	510	460	450	460	500
420	490	530	500	510	560	540	620
500	600	550	470	580	460	520	530
500	500	540	640	610	450	670	540

a. Determine the mean and standard deviation for this data.

b. The national mean for this test was 478. Assuming that these 40 students represent a random sample of the entire class, do you think Washington High's seniors did significantly better than the rest of the nation? Justify your response.

3.4 A set of scores on a physics exam are normally distributed with a mean of 72 and a standard deviation of 3. Four physics tests are found without names on them. The mean score of the four tests is 77. Is it reasonable to believe that these four tests came from this set of scores?

a. Formulate the null and alternative hypotheses for this situation.

b. Do you think that these four tests came from the physics class with a mean of 72? Explain your response.

3.5 An article in the *Daily News* reported that the mean height of women between the ages of 19 and 32 is 168 cm. In a letter to the editor, one reader insisted that this statement was untrue, arguing that the actual mean is less than 168 cm.

a. State the null and alternative hypotheses for the reader's claim.

b. To test the claim, the reader measured the heights of a random sample of 100 women between the ages of 19 and 32. The mean height in this sample was 164.5 cm, with a standard deviation of 16.2 cm. Find the z-score for the sample mean.

c. Using a 0.10 significance level, decide whether to reject, or fail to reject, the null hypothesis. Justify your reasoning.

d. What does your decision mean in this situation?

* * * * *

3.6 A liquid detergent company bottles their product in 1500-mL containers. After analyzing a random sample of 48 containers, a consumer group found a mean of 1488 mL. When the group called the company for an explanation, the customer relations department admitted that the product has a population standard deviation of 47.5 mL. Decide whether or not the detergent company is cheating its customers, with a 0.10 significance level. Justify your reasoning.

3.7 A breakfast cereal company claims that each box of its product contains at least 397 g of cereal. To test the manufacturer's claims, some students select a random sample of 40 boxes. The table below shows the mass of the cereal in each box, rounded to the nearest 1 g.

402	397	404	384	390	395	397	385	392	399
380	390	408	403	389	389	393	381	402	401
383	403	383	392	400	392	395	395	406	396
408	383	381	390	401	385	382	404	409	387

a. State the null and alternative hypotheses in this situation.

b. Determine the critical region for this hypothesis test at the 0.05 significance level.

c. Find the z-score for the sample statistic. **Note:** Use the standard deviation of the sample as an estimate of σ.

d. Decide whether to reject or fail to reject the null hypothesis. Justify your reasoning.

e. Explain what your decision means in this situation.

3.8 An elevator has a recommended capacity of 15 people and a maximum load limit of 1200 kg. The masses of the population that uses this elevator are normally distributed with a mean of 75 kg and a standard deviation of 10 kg.

a. What is the probability that a random sample of 15 people from this population will exceed the maximum load limit?

b. What is the probability that a random sample of 16 people will exceed the load limit?

* * * * * * * * * *

Summary Assessment

In a golden rectangle, the ratio of the measures of the longer side to the shorter side is the number $\left(1+\sqrt{5}\right)/2$, or about 1.618. The proportions of the golden rectangle are believed to be particularly pleasing to the human eye. For example, the outline of the Greek Parthenon resembles a golden rectangle, as does the face of each stone block in the Egyptian pyramids. In more modern times, the shapes of credit cards, driver's licenses, and the screens of many graphing calculators also are rough approximations of a golden rectangle.

The table below shows the length-to-width ratios of 30 beaded rectangles used to decorate Crow Indian leather goods:

1.706	1.706	1.704	1.178	1.176	1.175	1.916	1.502	1.919	1.912
1.504	1.499	1.890	1.701	1.894	1.704	1.887	1.698	1.880	1.942
1.159	2.237	1.751	2.242	2.232	1.754	1.748	2.066	2.075	2.070

From this data, does it appear that the Crow Indians also incorporated the golden ratio in their beaded designs? Use statistical analysis to support your belief. Assume that the data are normally distributed. **Note:** Use the standard deviation of the sample as an estimate of σ.

Module Summary

- In statistical analysis, there are two types of hypotheses. A **null hypothesis** (**H_0**) is a statement about one or more parameters. The **alternative hypothesis (**H_a**)** is the statement that must be true if the null hypothesis is false. The null hypothesis usually involves a claim of no difference or no relationship. In many situations, the null hypothesis and alternative hypothesis are negations of each other, but this is not necessarily the case.
- A **hypothesis test** may consist of the following steps.
  1. State null and alternative hypotheses about a parameter of a population.
  2. If the null hypothesis is true, predict what this implies about a sample of the population.
  3. Take a sample of the population and compare the results with your prediction.
  4. If the results are not consistent with the prediction, then you can conclude, with some level of certainty, that the null hypothesis is false and, therefore, reject it.
  5. If the results are consistent with the prediction, you fail to reject the null hypothesis. The failure to reject the null hypothesis does not guarantee that the null hypothesis is true, only that it might be true.
- Whenever a sample is taken of a population about which there is some uncertainty, it is possible that the sample is not representative of the population. Therefore, when performing a hypothesis test by sampling it is always possible to make an incorrect decision about the null hypothesis H_0. The two possible errors are rejecting a true null hypothesis and failing to reject a false null hypothesis.
- The curve that describes the shape of a normal distribution is the **normal curve**. The equation of the normal curve that models a normally distributed population depends on the mean and standard deviation of the population. The general equation for a normal curve, where μ represents the population mean and σ represents the population standard deviation, is:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2}$$

- The **standard deviation** σ is a measure of the variability, or spread, within a population. In a normally distributed population, the size of the standard deviation affects the height and width of the normal curve.

- The predictable variability of individual observations is summarized by the **68–95–99.7 rule**, which states that approximately 68% of the area between the x-axis and the normal curve is within 1 standard deviation of the mean, approximately 95% is within 2 standard deviations of the mean, and approximately 99.7% is within 3 standard deviations of the mean.
- Any value x from a normally distributed population with mean μ and standard deviation σ can be represented by a ***z*-score**. A z-score describes the number of standard deviations that the value is above or below the mean. The formula for determining a z-score is:

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

  The value in a table associated with this z-score represents the area between the x-axis and normal curve to the left of the value x.
- The **sampling distribution of sample means** is the distribution of sample means $\bar{x}$ of *all* possible samples of size n from a population.

  The mean of the sampling distribution of *all* sample means, denoted by $\mu_{\bar{x}}$, equals the population mean μ. The standard deviation of the sampling distribution, denoted by $\sigma_{\bar{x}}$, equals $\sigma/\sqrt{n}$, where σ is the population standard deviation and n is the sample size.
- The **central limit theorem** states that, even if the population from which samples are taken is not normally distributed, the sample means tend to be normally distributed. This approximation becomes more accurate as the sample size n increases. For $n \geq 30$, the distribution of sample means can be modeled reasonably well by a normal curve. This requirement is not necessary if the population from which samples are taken is normally distributed.
- Any value $\bar{x}$ from a normally distributed population with mean μ and standard deviation $\sigma_{\bar{x}}$ can be represented by a ***z*-score** as follows:

  $$z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$
- The **significance level** of a hypothesis test is an arbitrarily assigned probability that distinguishes a significant difference from a chance variation. It describes the maximum probability of making the error of rejecting a true null hypothesis.
- The **critical region** represents the set of all values that would lead a researcher to reject the null hypothesis. If the z-score of the individual observation (or sample mean) falls in the critical region, this indicates that the difference between the individual observation (or sample mean) and the population mean is significant, and not due to predictable variability. Therefore, the researcher should reject the null hypothesis.

Selected References

Moore, D. S. *Statistics: Concepts and Controversies*. New York: W. H. Freeman and Co., 1991.

Sincich, T. *Statistics by Example*. San Francisco: Dellen Publishing Co., 1987.

Sokal, R. R., and R. F. James. *Biometry*. San Francisco: W. H. Freeman and Co., 1969.

Travers, K. J., W. F. Stout, J. H. Swift, and J. Sextro. *Using Statistics*. Menlo Park, CA: Addison-Wesley, 1985.

Triola, M. F. *Elementary Statistics*. Reading, MA: Addison-Wesley, 1992.

Area under Normal Curve to Left of *Z*-Score										
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
–3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
–3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
–3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
–3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
–3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
–2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
–2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
–2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
–2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
–2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
–2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
–2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
–2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
–2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
–2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
–1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
–1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
–1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
–1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
–1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
–1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
–1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
–1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
–1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
–1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
–0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
–0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
–0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
–0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
–0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
–0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
–0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
–0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
–0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Area under Normal Curve to Left of *Z*-Score										
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998