Review for Exam 2

In my other class we check "conditions" instead of "assumptions". So when you see the word "condition" you know it means "assumption".

I like saying "Condition" better because it reminds me that I shouldn't do the test if the conditions aren't met.

0.1 Gathering and Describing Data Review

0.1.1 Gathering Data

Population: The set of all people, items, events, objects, etc. that are of interest.

Sample: A smaller part of the population on which we actually collect data.

Representative Samples: we want our samples to be similar to the population so we can use sample data to estimate population parameters

Random Samples: one of the best ways to get a representative sample is to randomly choose people to be in the sample.

Observational study: we simply observe and collect data. We don't try to influence the results or control any possible lurking variables. This means that we should **NOT** draw cause and effect conclusions.

Experiment: we try to control for as many possible lurking variables as possible. We use a treatment on the individuals and measure the response. One common method is to select two similar groups and impose a change on one group. We then measure the effect of the change or treatment. Because the only theoretical difference between the two groups is the treatment that we imposed, we should be able to draw cause and effect conclusions.

0.1.2 Types of Variables

Categorical Variable: A variable that places a case into one of several groups or categories.

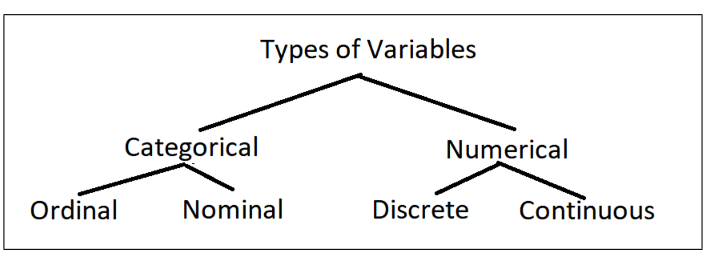
- A person's gender (male or female)
- Make of automobile (Honda, Ford, Dodge, etc.)

Numerical Variable: A variable that takes numerical values for which arithmetic operations such as adding and averaging make sense.

- Annual starting salary
- Gasoline mileage
- Height

Discrete Variable: A numerical variable X that can only take specific values.

Continuous Variable: A numerical variable X that can take any value in an interval of numbers.



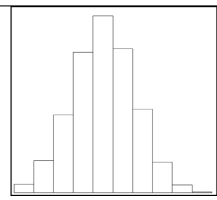
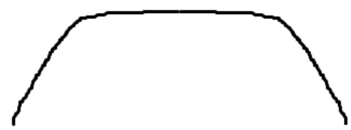
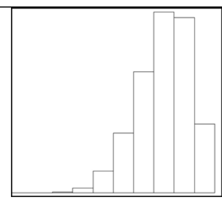
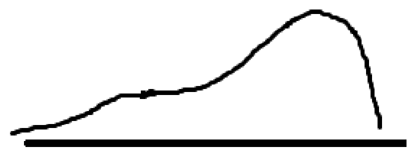
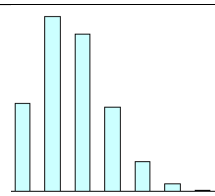
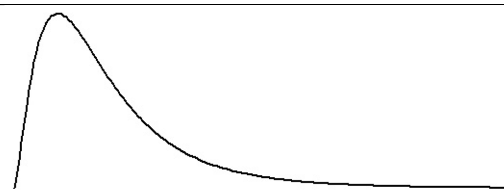
**We don't worry about the difference between ordinal and nominal variables in this class.

0.1.3 Shapes of Distributions

Histograms: Graph numerical data

- Divide the data into "classes" (also called groups or bins) of equal width.
- Display the counts or percentages of the observations that fall into each class.

Symmetry:

Symmetric Distributions		
Left Skewed Distributions		
Right Skewed Distributions		

Modes: the major peak of the distribution
(The mode is the value that occurs the most.)

Unimodal: one mode or peak

Bimodal: two modes or peaks

0.1.4 Describing Distributions with Numbers

0.1.4.1 Measure Center

Mean: the "average value"

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Median: the "middle value". Half the observations are smaller than the median and half the observations are bigger than the median. (When n is even, the median is the average of the two middle values.)

Resistant Measures:

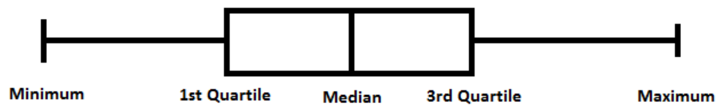
- The mean is not a resistant measure of the center. An outlier or even the observations in the tail of a skewed distribution can pull the mean towards the tail.
- The median is a resistant measure of the center. It is not as affected by extreme values.

0.1.4.2 Measure Spread

Quartiles: the quartiles divide the data into four equal parts; 25% of the data are in each part.

5 Number Summary: Minimum, 1st Quartile, Median, 3rd Quartile, Maximum

Boxplots: Use the 5 number summary to display the distribution.



Standard Deviation: measures spread by looking at how far the observations are from their mean.

0.1.4.3 Outliers

Outlier: A data value that doesn't seem to fit with the rest of the data.

1. Label each variable as numerical or categorical.

   If it is numerical, then break it down farther into continuous or discrete.

   (a) Your favorite brand of bread.

   (b) The percentage of whole wheat in your favorite bread.

   (c) The brand of your cell phone.

   (d) The battery life of your cell phone.

   (e) Whether or not a person has a credit card.

   (f) The number of books you own.

2. If we have a right skewed distribution, you would expect

   (a) the mean to be greater than the median.
   (b) the mean to be less than the median.
   (c) the mean to be equal to the median.

3. The mean is a resistant measure of central tendency.

   (a) True
   (b) False

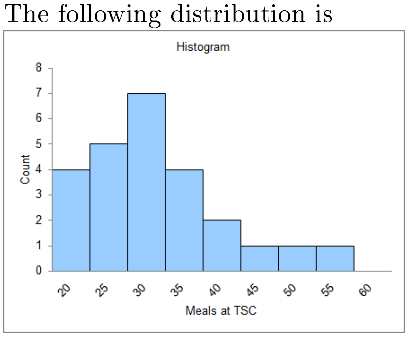
4. The median is a resistant measure to outliers and extreme values.

   (a) True
   (b) False

5. The following distribution is

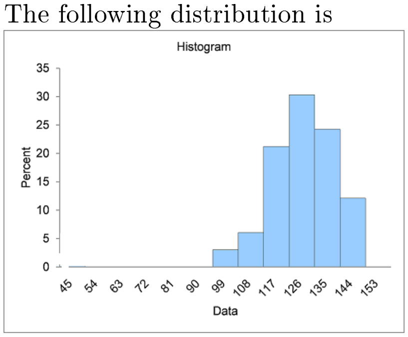
(a) left skew
(b) roughly symmetric
(c) right skew

6. The following distribution is



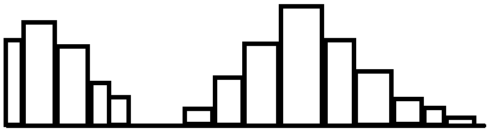
(a) left skew
(b) roughly symmetric
(c) right skew

7. The following distribution is



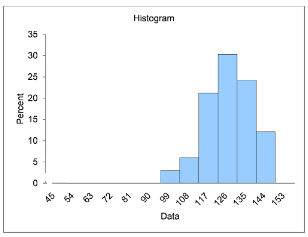
(a) unimodal
(b) bimodal
(c) three modes

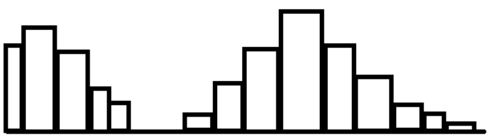
8. The following distribution is

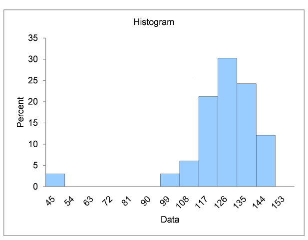


   (a) unimodal
   (b) bimodal
   (c) three modes

9. Which of the distributions has a potential outlier?

   (a) 

   (b) 

   (c) 

10. Label each variable as numerical or categorical.

    If it is numerical, then break it down farther into continuous or discrete.

    (a) The age of your puppy.

    (b) The publisher of your textbooks.

    (c) The weight of your backpack.

    (d) Whether or not you will graduate this year.

    (e) The type of credit card a person owns.

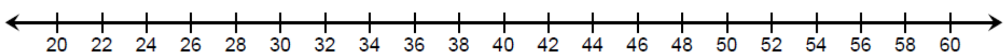
    (f) The credit limit on a person's credit card. Assume credit card companies only do $100 increments.

11. Here is a data set for the number of times 25 students ate at the TSC during the year.
23, 23, 23, 24, 26, 26, 26, 28, 29, 30, 31, 32, 32, 32, 34, 34, 36, 36, 37, 38, 40, 44, 46, 50, 56
Create a histogram of the data. Don't forget to label your axes.

class		count
lower boundary	upper boundary	
20	24	
25	29	
30	34	
35	39	
40	44	
45	49	
50	54	
55	59	

12. Use the summary information to create a boxplot for the meal data set.

Minimum	23
1st Quartile	26
Median	32
3rd Quartile	37
Maximum	56



13. The results of the test scores for a class are 81, 77, 99, 80, 95, 70, 43, 76, 95.

(a) What is the mean?

(b) What is the median?

14. One student didn't take the test, the data set is now 0, 43, 70, 76, 77, 80, 81, 95, 95, 99.

(a) Find the mean.

(b) Find the median.

(c) Compare your results to the results in problem 13. What changed more by adding the zero, the mean or the median?

0.2 Normal Distribution

Standard Normal Distribution: special normal distribution, mean $\mu = 0$, standard deviation $\sigma = 1$, use letter Z

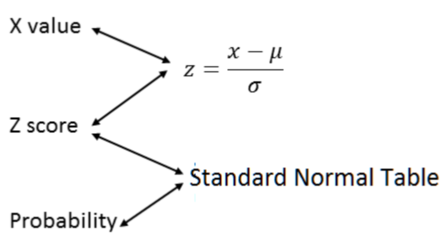
Empirical Rule: In the normal distribution with mean μ and standard deviation σ:

- Approximately 68% of the observations fall within 1σ of the mean μ.
- Approximately 95% of the observations fall within 2σ of μ.
- Approximately 99.7% of the observations fall within 3σ of μ. In the normal distribution with mean μ and standard deviation σ:

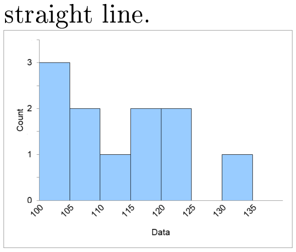
Z Score: The z score tells you how many standard deviations the x value is from the mean.
We call finding the z-score **standardizing** the random variable.

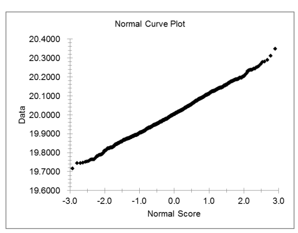
$$z = \frac{x - \mu}{\sigma}$$

Finding Probabilities: We can standardize and then use the standard normal table to find probabilities.
(But this is only accurate if the data is normally distributed.)



Normal Quantile Plots: If the data is approximately normal, the normal quantile plot will look like a fairly straight line.





15. ACT scores are reported on a scale from 1 to 36. The distribution of ACT scores for more than 1 million students in a recent high school graduating class was roughly Normal with mean $\mu = 20.8$ and standard deviation $\sigma = 4.8$.

    (a) Allen scores 27 on the ACT. What is the percentage of ACT scores that are higher than his score?

    (b) John wants to be in the top 10%. What score does he need?

16. SAT scores are reported on a scale from 400 to 1600. The distribution of SAT scores for 1.4 million students in the same graduating class was roughly Normal with mean $\mu = 1026$ and standard deviation $\sigma = 209$.

    (a) Reports on a student's SAT usually give the percentile as well as the actual score. The percentile is just the cumulative proportion stated as a percent: the percent of all scores that were lower than this one. Jessica scores 880 on the SAT. What is her percentile?

    (b) Melissa is hoping to qualify for a scholarship. Those are only awarded to applicants who have a SAT score among the top 5% of all SAT scores. What is the lowest score that she can get that will be good enough to qualify for a scholarship?

    (c) We decide to randomly select one student. What is the probability that his score is less than 1200?

17. Compare a SAT with an ACT score: Wendy scores 1350 on the SAT. Jeremy scores 25 on the ACT. Assuming that both tests measure the same thing, who has the higher score – Wendy or Jeremy? Report the z-scores for both students.

18. Use the 68-95-99.7 rule (empirical rule) to describe each data set.

    (a) IQ test scores are supposed to be normally distributed with a mean of 100 and a standard deviation of 15.

    (b) Suppose teenage boys in Utah are normally distributed with a mean of 66.5 inches and a standard deviation of 3.5 inches.

    (c) Egg lengths are supposed to be normally distributed with a mean of 6 cm and a standard deviation of 1.4 cm.

19. The temperature at any random location in a kiln used for manufacturing bricks is normally distributed with a mean of 1000ºF and a standard deviation of 50ºF.

    (a) What is the z-score that relates to 1075ºF?

    (b) A z-score of -2.75 relates to a temperature of ______ºF?

    (c) If bricks are fired at a temperature above 1125ºF, they will crack and must be discarded. If the bricks are placed randomly throughout the kiln, what is the percentage of bricks that crack during the firing process?

    (d) When glazed bricks are put in the oven, if the temperature is below 900ºF, they will discolor. If the bricks are placed randomly throughout the kiln, what percentage of glazed bricks will discolor?

    (e) When completely filled, bricks at the top 10% hottest locations will be exposed to temperatures of ________ºF (and higher).

0.3 Sampling Distributions

0.3.1 Parameters versus Statistics

Population: The set of *all* people, items, events, objects, etc. that are of interest.

Sample: a subset of the population on which we collect data.

Population Parameter: a number that describes a characteristic of the population.

Parameter: property of an underlying probability distribution

Sample Statistic: a number that describes a characteristic of the population.
We use the sample statistic to make inferences about the population parameter.

Because the value of the sample statistic changes based on which sample we choose, we call it a random variable. So things like the sample mean $\overline{X}$, the sample standard deviation S, the sample proportion $\hat{p}$, etc. are all random variables.

0.3.2 All Possible Samples

From any population, there are many different samples of size n that can be chosen.

Sampling Distribution

- Since there are different samples that can be chosen, any sample statistic, (i.e. sample mean, sample mode, sample median, sample first quartile, sample standard deviation, sample range, sample maximum value, etc.), will be different based on which sample is chosen.
- Each statistic will be a random variable because its value changes from sample to sample.
  - That means that each statistic also has its own distribution, called the sampling distribution.
  - The idea that the value of the statistic changes from sample to sample is called **sampling variability.**

0.3.3 Estimating

We use our sample statistics to estimate population parameters.

- $\overline{x}$ is the best estimate of μ
- $\hat{p}$ is the best estimate of p
- s is the best estimate of σ
- etc.

The standard error tells us how close to the true population parameter our estimate is likely to be.

0.4 Symbols for Means and Standard Deviations

0.4.1 Individual Values in a Population

μ**:** the mean for the ***original population of individual values***

σ**:** the standard deviation for the ***original population of individual values***

p**:** the proportion for the ***original population of individual values***

0.4.2 Individual Values in a Sample

$\overline{x}$**:** the sample mean of the individual observations in a single sample

s**:** the sample standard deviation of the individual observations in a single sample

$\hat{p}$**:** the sample proportion of the individual observations in a single sample

0.4.3 Sampling Distribution of Sample Mean

$\mu_{\overline{x}}$**:** the theoretical mean for the population of ***all possible sample means***

$\sigma_{\overline{x}}$**:** the theoretical standard deviation for the population of ***all possible sample means***

$SE_{\overline{x}}$**:** (standard error of $\overline{x}$) the sample estimate of the standard deviation for ***all possible sample means***

0.4.4 Sampling Distribution of Sample Proportion

$\mu_{\hat{p}}$**:** the theoretical mean for the population of ***all possible sample proportions***

$\sigma_{\hat{p}}$**:** the theoretical standard deviation for the population of ***all possible sample proportions***

$SE_{\hat{p}}$**:** (standard error of $\hat{p}$) the sample estimate of the standard deviation for ***all possible sample proportions***

0.5 Sampling Distribution Formulas

The Theoretical Population of All Possible Sample Means

- **Shape:** Normal if population is normal or our sample size is at least 30
- **Center:** $\mu_{\overline{x}} = \mu$ (the theoretical mean of all the possible sample means is equal to the original population mean)
- **Spread:** $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$ (the theoretical standard deviation of all the possible sample means is equal to the original population standard deviation divided by the square root of n)

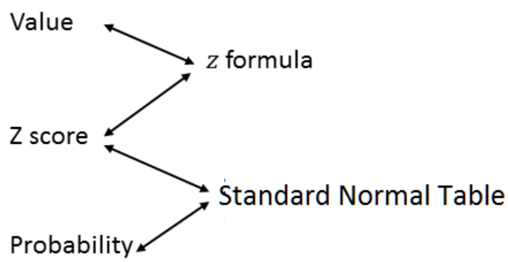
The Theoretical Population of All Possible Sample Proportions

- **Shape:** Normal if $np \geq 5$ and $n\left(1-p\right) \geq 5$
- **Mean:** $\mu_{\hat{p}} = p$ (the theoretical mean of all possible sample proportions is equal to the original population proportion)
- **Standard Deviation:** $\sigma_{\hat{p}} = \sqrt{\dfrac{p\left(1-p\right)}{n}}$ (the theoretical standard deviation of all the possible sample proportions...)
- **Standard Error:** $S.E.\left(\hat{p}\right) = \sqrt{\dfrac{\hat{p}\left(1-\hat{p}\right)}{n}}$ (our estimate of $\sigma_{\hat{p}}$ using the sample proportion)

The Theoretical Population of All Possible Sample Counts

- **Shape:** Normal if $np \geq 5$ and $n\left(1-p\right) \geq 5$
- **Mean:** $\mu_X = np$
- **Standard Deviation:** $\sigma_X = \sqrt{np\left(1-p\right)}$

0.6 Find Probabilities



Probabilities for Individual Values

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

Assumption: normal population

Probabilities for Sample Means

$$z = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$$

Assumption: normal population or $n \geq 30$

Probabilities for Sample Count

Don't forget continuity correction (add/subtract .5 as appropriate).

**Switch proportions to counts.

$$z = \frac{X - np}{\sqrt{np\,(1 - p)}}$$

Assumption: $np \geq 5$ and $n\,(1 - p) \geq 5$

20. All the books published by a certain publisher have a mean of 256 pages and a standard deviation of 23 pages. A random sample of books by the publisher have a mean of 243 pages and a standard deviation of 19 pages.

    (a) Summarize the given information with appropriate symbols.

    (b) Label each number as a parameter or statistic.

21. In a controversial election district, **73%** of registered voters are Democrats. A random survey of 500 voters had **68%** Democrats. Are the bold numbers parameters or statistics?

    (a) Both are statistics.
    (b) 73% is a parameter and 68% is a statistic.
    (c) 73% is a statistic and 68% is a parameter.
    (d) Both are parameters.

22. SAT scores are reported on a scale from 400 to 1600. The distribution of SAT scores for 1.4 million students in the same graduating class was roughly Normal with mean $\mu = 1026$ and standard deviation $\sigma = 209$. Suppose we take a random sample of 13 seniors.

    (a) What is the mean of all the possible sample means?

    (b) What is the standard deviation of the sampling distribution of the sample means?

    (c) What is the shape of the distribution of all the possible sample means?

    (d) What is the probability that the sample mean is less than 1200?

    (e) We decide to randomly select one student. What is the probability that his score is less than 1200?

    (f) What is the probability that the sample mean is between 1100 and 1200?

23. Packages of sugar bags for Sweeter Sugar Inc. are supposed to have an average weight of 16 ounces and a standard deviation of 0.3 ounces and are normally distributed. The company wants to find a sample of 15 bags and check the average weight.

(a) Let X be the the weight of one bag of sugar. Why is X a random variable?

(b) Why is $\overline{X}$ a random variable?

(c) What will the shape of the distribution of the population of all possible $\overline{X}$'s be?

(d) What is $\mu_{\overline{X}}$?

(e) What is $\sigma_{\overline{X}}$?

(f) What is the probability that 15 randomly selected packages will have a average weight in excess of 16.1 ounces?

(g) What is the probability that 1 randomly selected bag will have a weight between 15 and 17 ounces?

24. The population lifetime of a certain battery has a mean of $\mu = 153$ hours and a standard deviation of $\sigma = 36$ hours. The distribution is right skewed. We decide to take a sample of 10 batteries and look at the sample mean.

(a) What is the mean of the sampling distribution of $\overline{X}$?

(b) What is the standard deviation of the sampling distribution of $\overline{X}$?

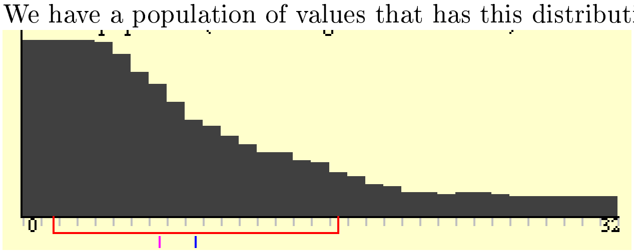
(c) What is the shape of the sampling distribution of $\overline{X}$?

(d) Find the probability that the sample mean is less than 140 hours.

25. Which distribution will have a smaller standard deviation:

   (a) a population of individual values
   (b) the sampling distribution of the sample mean for a sample of size $n = 2$?

26. We have a population of values that has this distribution:



What will the shape of the sampling distribution of the sample mean for $n = 3$ look like?

   (a) normal
   (b) like the original population
   (c) still right skewed, but a little more normal than the original population is

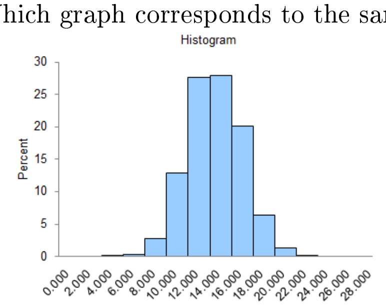
27. Which sample size will give the smallest standard deviation for the sampling distribution of the sample mean?

   (a) $n = 100$.
   (b) $n = 500$.
   (c) $n = 1000$.
   (d) It depends on the distribution, i.e., whether it is symmetric or skewed.

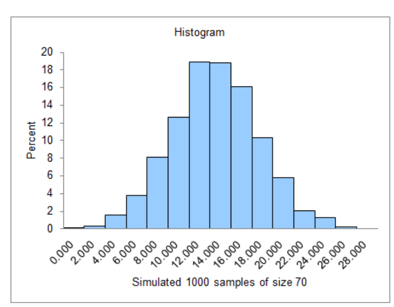
28. The true percentage of red M&Ms is 15%. If six students took a sample of M&Ms and recorded the percentage of red M&Ms, which set of values are they most likely to obtain?

   (a) 15%, 15%, 15%, 15%, 15%
   (b) 3%, 90%, 40%, 86%, 15%
   (c) 14%, 13%, 15%, 14%, 16%
   (d) Any of the above

29. The graphs below are the sampling distributions for the percentage of red M&Ms for sample sizes of 70 and 200. Which graph corresponds to the sample size of 200?



A.



B.

30. We know that percentage of boy births in the US is 51.2%. At which hospital are we most likely to observe a week with 83% boy births?

    (a) a small hospital with five births per week

    (b) a medium hospital with fifty births per week

    (c) a large hospital with a hundred births per week

31. It is estimated that 75% of all young adults between the ages of 18-35 do not have a landline in their homes and only use a cell phone at home. We decide to take a random sample of 100 young adults and find the sample proportion $\hat{p}$.

    (a) What is the mean of the sampling distribution of $\hat{p}$?

    (b) What is the standard deviation of the sampling distribution of $\hat{p}$?

    (c) What is the shape of the sampling distribution of $\hat{p}$?

32. We believe that the proportion of all teenagers who listen to streamed music is 30%. Let's take a sample of 1000 teenagers.

    (a) What is n? What is p?

    (b) What is the mean of the sampling distribution of $\hat{p}$?

    (c) What is the standard deviation of the sampling distribution of $\hat{p}$?

    (d) Can you use the normal distribution as an approximation for this problem?

0.7 Confidence Intervals

0.7.1 Find Confidence Interval

0.7.1.1 Z Confidence Interval for Population Mean μ (σ is known)

$\overline{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

condition: normal population or $n \geq 30$

0.7.1.2 T Confidence Interval for Population Mean μ (σ is unknown)

$\overline{x} \pm t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n}}$

Degrees of freedom: $n-1$

condition: normal population or $n \geq 30$

0.7.2 Find sample size for desired margin of error

sample size for Z confidence interval for mean: $n = \left(\dfrac{z_{\alpha/2}\sigma}{m}\right)^2$

***Note that the review was written for another class so the key uses the notation z^* instead of $z_{\alpha/2}$ and t^* instead of $t_{\alpha/2,n-1}$.

Multiple Choice

33. Determine whether each of the following statements is true or false.
A) The margin of error for a 95% confidence interval for the mean μ increases as the sample size increases.
B) The margin of error for a confidence interval for the mean μ increases as the confidence level decreases.
C) The margin of error for a 95% confidence interval for the mean μ decreases as the population standard deviation decreases.

34. A small New England college requires the Math SAT for admission, and the population standard deviation is found to be 60. The formula for a 95% confidence interval yields the interval 634.12 to 645.88. Determine whether each of the following statements is true or false.
A) If we repeated this procedure many, many times, only 5% of all the possible confidence intervals would NOT include the mean Math SAT score of the population of all students at this college.
B) The probability that the population mean will fall between 634.12 and 645.88 is 0.95.
C) If we repeated this procedure many, many times, $\overline{x}$ would fall between 634.12 and 645.88 about 95% of the time.

35. A sprinkler system is being installed in a large office complex. Based on a series of test runs, a 99% confidence interval for μ, the average activation time of the sprinkler system (in seconds), is found to be (22, 28). Determine whether each of the following statements is true or false.
A) The 99% confidence level implies that $P(22 < \mu < 28) = 0.99$.
B) The 99% confidence level implies that 99% of the sample means ($\overline{x}$) obtained from repeated sampling would fall between 22 and 28.

36. Central Middle School has calculated a 95% confidence interval for the mean height μ of eleven- year-old boys at their school and found it to be 56 $\pm$ 2 inches.

    (a) Determine whether each of the following statements is true or false.
    A) There is a 95% probability that μ is between 54 and 58.
    B) There is a 95% probability that the true mean is 56.
    C) If we took many additional random samples of the same size and from each computed a 95% confidence interval for μ, approximately 95% of these intervals would contain μ.
    D) If we took many additional random samples of the same size and from each computed a 95% confidence interval for μ, approximately 95% of the time μ would fall between 54 and 58.

    (b) Which of the following could be the 90% confidence interval based on the same data?
    A) 56 $\pm$ 1
    B) 56 $\pm$ 2
    C) 56 $\pm$ 3
    D) Without knowing the sample size, any of the above answers could be the 90% confidence interval.

37. An agricultural researcher plants 25 plots with a new variety of yellow corn. Assume that the yield per acre for the new variety of yellow corn follows a Normal distribution with an unknown mean of μ and a standard deviation of $\sigma = 10$ bushels per acre.

   (a) If the average yield for these 25 plots is $\overline{x} = 150$ bushels per acre, what is a 90% confidence interval for μ?
   A) 150 ± 0.784
   B) 150 ± 2.00
   C) 150 ± 3.29
   D) 150 ± 3.92

   (b) Which of the following would produce a confidence interval with a smaller margin of error than the 90% confidence interval?
   A) Plant only five plots rather than 25, because five are easier to manage and control.
   B) Plant 10 plots rather than 25, because a smaller sample size will result in a smaller margin of error.
   C) Plant 100 plots rather than 25, because a larger sample size will result in a smaller margin of error.
   D) Compute a 99% confidence interval rather than a 90% confidence interval, because a higher confidence level will result in a smaller margin of error.

38. To assess the accuracy of a laboratory scale, a standard weight that is supposed to weigh exactly 1 gram is repeatedly weighed a total of n times and the mean $\overline{x}$ is computed. Suppose the scale readings are Normally distributed with an unknown mean of μ and a standard deviation of $\sigma = 0.01$ g. How large should n be so that a 95% confidence interval for μ has a margin of error no larger than $\pm$ 0.0001?
A) $n = 100$
B)$n = 196$
C) $n = 10000$
D)$n = 38416$

39. The heights of a simple random sample of 400 male high school sophomores in a Midwestern state are measured. The sample mean is $\overline{x} = 66.2$ inches. Suppose that the heights of male high school sophomores follow a Normal distribution with a standard deviation of $\sigma = 4.1$ inches.

   (a) What is a 95% confidence interval for μ?
   A) (58.16, 74.24)
   B) (59.46, 72.94)
   C) (65.80, 66.60)
   D) (65.86, 66.54)

   (b) Suppose the heights of a simple random sample of 100 male sophomores were measured rather than 400. Which of the following statements is true?
   A) The margin of error for the 95% confidence interval would increase.
   B) The margin of error for the 95% confidence interval would decrease.
   C) The margin of error for the 95% confidence interval would stay the same because the level of confidence has not changed.

40. Suppose we wish to calculate a 90% confidence interval for the average amount spent on books by freshmen in their first year at a major university. The interval is to have a margin of error of \$2. Assume that the amount spent on books by freshmen has a Normal distribution with a standard deviation of $\sigma = \$30$. How many observations are required to achieve this margin of error?
A) 25
B) 608
C) 609
D) 865

41. Battery packs in radio-controlled racing cars need to be able to last pretty long. The distribution of the lifetimes of battery packs made by Lectric Co. is slightly left-skewed. Assume that the standard deviation of the lifetime distribution is $\sigma = 2.5$ hours. A simple random sample of 75 battery packs results in a mean of $\overline{x} = 29.6$ hours.

   (a) What is a 90% confidence interval for μ, the true average lifetime of the battery packs made by Lectric Co.?
   A) (29.13, 30.07)
   B) (29.03, 30.17)
   C) (28.86, 30.34)
   D) The confidence interval cannot be calculated because the population distribution is not Normal.

   (b) Which statement is true:
   A) If a 95% confidence interval had been calculated, the margin of error would have been larger.
   B) If many more samples of 75 battery packs were taken, 90% of the resulting confidence intervals would have a sample mean between 29.13 and 30.07.
   C) If the sample size had been 150 and not 75, the margin of error would have been larger.

42. A nationally distributed college newspaper conducts a survey among students nationwide every year. This year, responses from a simple random sample of 204 college students to the question "About how many CDs do you own?" resulted a 95% confidence interval for the mean number of CDs owned by all college students. The interval is (71.8, 73.8).

   (a) Answer each of the following questions with yes, no, or can't tell.
   A) Does the sample mean lie in the 95% confidence interval?
   B) Does the population mean lie in the 95% confidence interval?
   C) If we were to use a 92% confidence level, would the confidence interval from the same data produce an interval wider than the 95% confidence interval?
   D) With a smaller sample size, all other things being the same, would the 95% confidence interval be wider?

   (b) Which of the following interpretations of the interval is correct?
   A) 95% of all students own between 71.8 and 73.8 CDs
   B) the probability that the population mean is between 71.8 and 73.8 CDs is 95%
   C) we are 95% confident that the sample mean is between 71.8 and 73.8
   D) we are 95% confident that the population mean number of CDs is between 71.8 and 73.8

Show Your Work

43. A total of 114 male athletes from Canadian sports centers were surveyed. The average calorie intake was 3077 kcal/day with a sample standard deviation of 987.

   (a) Summarize the given information with appropriate symbols.

   (b) Which confidence interval should you use?

   (c) Check conditions.

   (d) Find the **99%** confidence interval for the population mean calorie intake μ.

   (e) What is the margin of error?

   (f) Interpret your result.

44. A nationally distributed college newspaper conducts a survey among students nationwide every year. This year, responses from a simple random sample of 204 college students to the question “About how many CDs do you own?” resulted in a sample mean of $\overline{x} = 72.8$. Based on data from previous years, the editors of the newspaper will assume that $\sigma = 7.2$. They want to find a confidence interval.

(a) Which formula should you use?

(b) Check conditions.

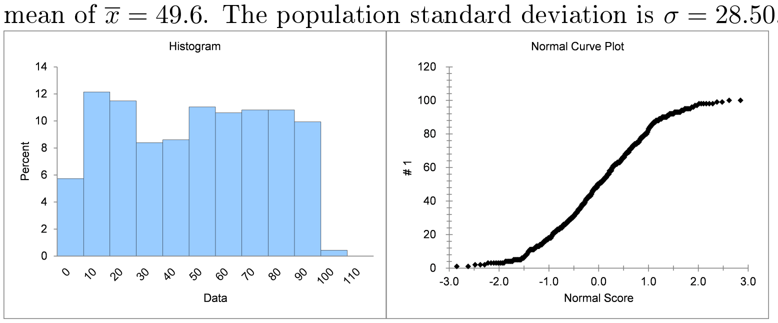
(c) Find the 97% confidence interval for the population mean number of CDs μ.

(d) What is the margin of error?

(e) Interpret your confidence interval.

(f) What does the 97% confidence level mean?

45. You want to find a 90% confidence interval for the population mean μ. The sample size is $n = 21$ with a sample mean of $\overline{x} = 49.6$. The population standard deviation is $\sigma = 28.50$.



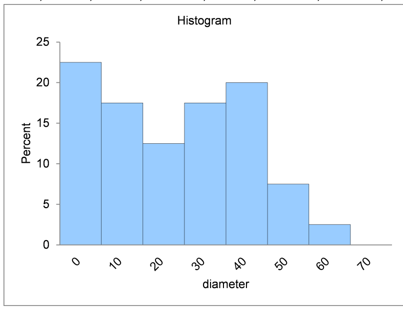
(a) Which formula should you use?

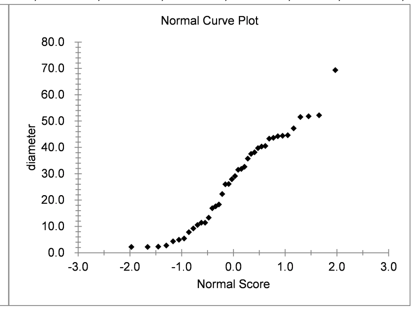
(b) Check conditions.

(c) Find the 90% confidence interval for the population mean μ.

(d) ***NOTE!!! The plot above is called a NORMAL CURVE PLOT or a NORMAL QUANTILE PLOT. A lot of students have thought that if the problem on the exam says **normal** quantile plot that the data is normal. That is not true. It is just the name of the plot. You have to look at the plot to see if it is close to a straight line.

46. We have a data set of the diameter of trees in Georgia. Here is a sample of 40 trees.
Data Set: 10.5, 13.3, 26, 18.3, 52.2, 9.2, 26.1, 17.6, 40.5, 31.8, 47.2, 11.4, 2.7, 69.3, 44.4, 16.9, 35.7, 5.4, 44.2, 2.2, 4.3, 7.8, 38.1, 2.2, 11.4, 51.5, 4.9, 39.7, 32.6, 51.8, 43.6, 2.3, 44.6, 31.5, 40.3, 22.3, 43.3, 37.5, 29.1, 27.9





The sample mean is 27.69 and the standard deviation is 17.706.

(a) Summarize the given information with appropriate symbols.

(b) Which confidence interval formula should you use?

(c) Check conditions.

(d) Find the 95% confidence interval for the population mean tree diameter μ.

(e) Interpret your result.

0.8 Review for Hypothesis Tests

0.8.1 Notation

	size	mean	standard deviation	variances
population parameter		μ	σ	σ^2
sample statistic	n	$\overline{x}$	s	s^2

0.8.2 What I expect to see on each hypothesis test problem:

- Which test should you use?
- Choose significance level.
- Check the conditions (assumptions).
- What are the hypotheses?
- What is the test statistic?
- What is the p-value?
- Do you reject or fail to reject the null hypothesis?
- Interpret your conclusion.

0.8.3 One Sample Z Test (know population standard deviation σ)

Condition: Normal Population or $n \geq 30$

$H_0 : \mu = \mu_0$

$z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$

0.8.3.1 Doing a Z test with your calculator

I had you do the Z test by hand but now you can just put it in your calculator like you do the T test if you prefer.

0.8.4 One Sample T Test (know sample standard deviation s)

Condition: Normal Population or $n \geq 30$

$H_0 : \mu = \mu_0$

$t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$

$df = n - 1$

0.8.5 Miscellaneous

Null Hypothesis: H_0, The status quo or the statement being tested.

Alternative Hypothesis: H_a, The opposite of the null hypothesis.

Test Statistic: Either $z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$ or $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$ depending on the test.

Level of Significance: The significance level α is the amount of evidence against the null hypothesis we require to reject H_0.

P-value: The p-value is the area in the tail(s).

The P-value is the probability of observing our test statistic or something more extreme if our null hypothesis is true.

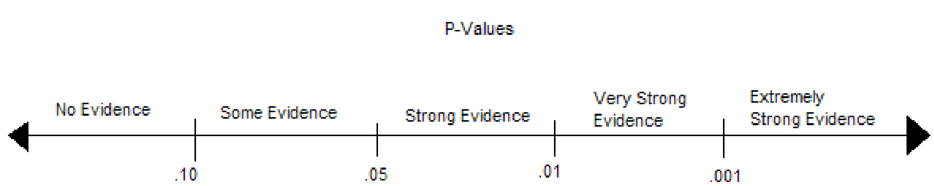
The smaller the p-value, the more evidence you have against the null hypothesis!

0.8.6 Errors

Type I Error: When we reject a true null hypothesis

Type II Error: When we fail to reject a false null hypothesis

0.8.7 Significance and P-values



- The level of significance α tells us how much evidence we will require to reject the null hypothesis.
- The p-value tells us how much evidence we actually found against the null hypothesis.
- Statistical significance is different from practical significance.

47. The cost of a one-carat VS2 clarity, H color diamond from Diamond Source USA is $5600.
A Midwestern jeweler makes calls to contacts in the diamond district of New York City to see whether the mean price of diamonds there differs from $5600. He calls 35 jewelers and the results are a sample mean of $5835. He believes the population standard deviation is $520.
Conduct a test to see if the mean diamond price in New York City is different from $5600.

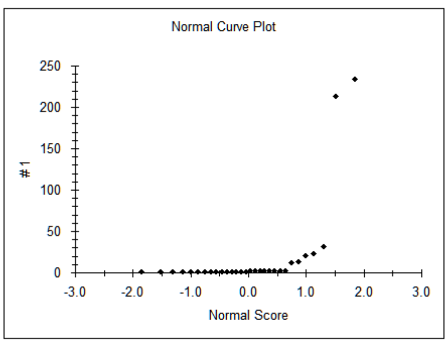
48. A hamburger restaurant wants to form a partnership with a potato company. The potatoes come in supposedly 20 pound bags. Of course the hamburger restaurant doesn't want to buy underweight bags of potatoes. Before they sign a contract, they weigh a sample of 40 bags and find an average weight of 19.7 pounds with a standard deviation of 0.1 pounds. Conduct a test to see if the population mean of the bag weights will be less than 20 pounds. **Interpret the p-value.**

49. Nike, thinks it has found a way to make running shoes that last longer than the current average of 750 miles. A sample of 35 runners are selected to wear the new design and the shoes last an average of 815 miles. Assume that the population standard deviation is 25 miles. Conduct a test to see if the average distance the new design will last is greater than 750 miles. **Interpret the p-value.**

50. The Employment and Training Administration reported the US mean unemployment insurance benefit of \$238 per week. A researcher in Virginia wonders if they have a different benefit than the national average.
He takes a sample of 85 individuals in Virginia and finds the sample mean weekly unemployment benefit was \$231 with a sample standard deviation of \$80.
Conduct a hypothesis test to see if Virginia has a different mean benefit than the national average.

51. Your boss wants you to run a hypothesis test to see if $\mu < 87$ for a data set with sample size 20. You found the normal curve plot below.

(a) What will you tell your boss?



(a) ***NOTE!!! The plot above is called a NORMAL CURVE PLOT or a NORMAL QUANTILE PLOT. A lot of students have thought that if the problem on the exam says **normal** quantile plot that the data is normal. That is not true. It is just the name of the plot. You have to look at the plot to see if it is close to a straight line.

More Review for Hypothesis Tests

52. A statistics class also took a sample of 30 Chips Ahoy Cookies. They claim to have 1000 chocolate chips per bag which means an average of 33 chocolate chips per cookie. We want to know if there are less than an average of 33 chocolate chips per cookie.

    (a) What are the null and alternative hypotheses in words?

    (b) The sample mean is $\overline{x} = 24.83$ and the sample standard deviation is $s = 3.65$. The p-value is 2.7×10^{-13}. Interpret the p-value.

53. A hamburger restaurant wants to form a partnership with a potato company.
    The potatoes come in supposedly 20 pound bags. Of course the hamburger restaurant doesn't want to buy underweight bags of potatoes.
    Before they sign a contract, they weigh a sample of 40 bags and find an average weight of 19.7 pounds with a standard deviation of 0.1 pounds.
    They conduct a test to see if the population mean of the bag weights will be less than 20 pounds.
    Their p-value is 1.35×10^{-21}.

    (a) The small p-value means we have _________________ significance.

    i. practical
    ii. statistical

    (b) Which interpretation of the p-value is correct?

    i. The probability that the mean weight is 20 pounds is 1.35×10^{-21}.
    ii. The probability that the mean weight is less than 20 pounds is 1.35×10^{-21}.
    iii. If the population mean really is 20 pounds, then the probability of getting a sample mean of $\overline{x} = 19.7$ or less, is 1.35×10^{-21}.

54. What does the level of significance $\alpha = .05$ tell us?

55. What is the p-value?

    (a) The probability that the null hypothesis is true.
    (b) The probability that the alternative hypothesis is true.
    (c) The probability that the null hypothesis is false.
    (d) The probability of observing our sample data, or something more extreme, if the null hypothesis is true.

56. In the cookie problem, $H_0 : \mu = 33$ versus $H_A : \mu < 33$, we got a p-value of essentially 0 (software tells us it was 1.95×10^{-32}). What is an appropriate conclusion?

   (a) We have proved that the population mean number of chocolate chips is 33.
   (b) We have proved that the population mean number of chocolate chips is less than 33.
   (c) We found extremely strong evidence that the population mean number of chocolate chips is 33.
   (d) We found extremely strong evidence that the population mean number of chocolate chips is less than 33.

57. A KFC manager conducts a hypothesis test with $H_0 : \mu = 1$ versus $H_A : \mu > 1$. He decided ahead of time to use a significance level of $\alpha = .01$. His p-value is .0367.

   (a) Did he reject or fail to reject the null hypothesis?

   (b) Is it ethical for him to change his significance level after he sees the results?

   (c) If he had chosen a significance level of .05 instead of $\alpha = .01$ ahead of time, would he reject or fail to reject the null hypothesis?

58. Suppose unbeknownst to our manager, managers at several different KFCs did the same experiment. The different p-values that they found are shown below. Which manager found the strongest evidence that the $\mu > 1$.

   (a) .023
   (b) .564
   (c) .001
   (d) .00034

59. Statistically significant means that it is very unlikely that we would get such an extreme sample statistic if the null hypothesis were true.
We say that our results are statistically significant if we get a ____________ p-value.

    (a) big
    (b) small

60. We conducted a hypothesis test to determine if a dice was unfair.

    (a) What would be the null hypothesis (in words)?

    (b) What would be the alternative hypothesis (in words)?

    (c) What would be a Type I error?

    (d) What would be a Type II error?

61. The ____ the P-value, the stronger the evidence against the null hypothesis provided by the data.
A) larger
B) smaller

62. When performing significance tests at level α, if the P-value obtained is _____ α, the null hypothesis will not be rejected.
A) larger than
B) smaller than
C) the same as
D) None of the above.

63. When performing significance tests at level α, if the P-value obtained is _____ α, the null hypothesis will be rejected.
A) larger than
B) smaller than
C) the same as
D) None of the above.

64. If the P-value is as small as or smaller than a specified value of α, then the data are _______ significant.
A) practically
B) statistically
C) absolutely

65. A test of significance for a null hypothesis has been conducted and the P-value determined. Which of the following statements about a P-value is (are) TRUE?
A) The P-value is the probability that the null hypothesis is false.
B) The P-value is the probability that the alternative hypothesis is true.
C) The P-value is the probability that the null hypothesis is rejected even if that hypothesis is actually true.
D) The P-value tells us the strength of the evidence against the null hypothesis.
E)The larger the P-value the stronger the evidence against the null hypothesis.

66. If a statistical significance is found when performing a test of significance, then practical significance will also be found.
A) True
B) False

67. If you reject the null hypothesis when in fact the null hypothesis is true it is called ___________.
A) a Type I error
B) a Type II error

68. If you accept the null hypothesis when in fact the alternative hypothesis is true is called ___________.
A) a Type I error.
B) a Type II error.

69. A small company consists of 25 employees. As a service to the employees, the company arranges for each of the employees to have a complete physical exam for free. Among other things, the weight of each employee is measured. The mean weight is found to be 165 pounds. The standard deviation of the weight measurements is 20 pounds. It is believed that a mean weight of 160 pounds would be expected for this group. To see if there is evidence that the mean weight of the population of all employees of the company is significantly larger than 160, the hypotheses $H_0 : \mu = 160$ versus $H_a : \mu > 160$ are tested. You obtain a P-value of about 0.106. Determine which statement is true for a significance level of $\alpha = .05$.

(a) You have proved H_0 is true.
(b) You have proved H_a is true.
(c) You have proved H_0 is false.
(d) You failed to find evidence against H_0.

70. Ten years ago, at a small high school in Alabama, the mean Math SAT score of all high school students who took the exam was 490 with a standard deviation of 80. This year, the Math SAT scores of a random sample of 25 students who took the exam are obtained. The mean score of these 25 students is $\overline{x} = 525$. We assume the population standard deviation continues to be $\sigma = 80$. To determine if there is evidence that the scores in the district have improved, the hypotheses $H_0 : \mu = 490$ versus $H_a : \mu > 490$ are tested. The P-value is found to be 0.014. Use the significance level of $\alpha = .05$. What conclusion can we draw?

(a) We have proved that the mean score is 490.
(b) We have proved that the mean score is greater than 490.
(c) We found evidence that the mean score is greater than 490.
(d) We didn't find any evidence that the mean score is greater than 490.

71. A study was conducted to determine the effect of the addition of a particular supplement to a well-known low-density lipoprotein (LDL) cholesterol lowering drug. 5000 randomly selected patients were studied on this combination of drugs over a one-year period. An increased reduction of 5 mg/dL of LDL cholesterol was achieved with the combination of drugs relative to the well-known drug alone. The reduction resulted in a P-value of 0.001.
A) This P-value result shows that the combination drug should be prescribed widely.
B) Clearly, this study has demonstrated the effectiveness of this new treatment.
C) The result should be viewed with caution because one year is not a long time to study such an important matter as cholesterol.
D) A sample of this size can detect very small differences and hence care should be taken to determine if a reduction of this amount has any practical significance with respect to LDL cholesterol reduction.

72. What is power?
   (a) The probability of rejecting a true null hypothesis.
   (b) The probability of rejecting a false null hypothesis.
   (c) The probability of accepting a true null hypothesis.
   (d) The probability of accepting a false null hypothesis.

73. Higher power is a good thing. How can we increase the power and lower the probability of making an error?

74. Read this: Does taking ginkgo tablets twice a day provide significant improvement in mental performance? To investigate this issue, a researcher conducted a study with 150 adult subjects who took ginkgo tablets twice a day for a period of six months.

At the end of the study, 200 different variables related to the mental performance of the subjects were measured on each subject and the means compared to known means for these variables in the population of all adults.

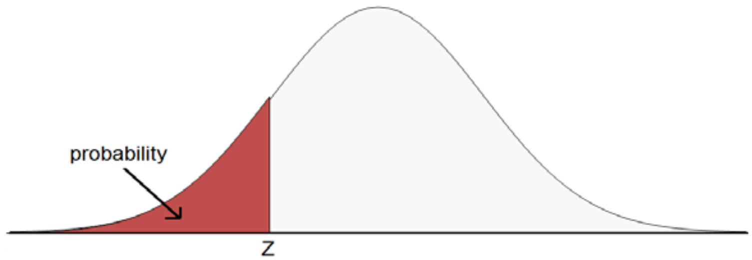
(So 200 hypothesis tests were conducted.)

Nine of these variables were significantly better (in the sense of statistical significance) at the 5% level for the group taking the ginkgo tablets as compared to the population as a whole.

Do you think this means that we can say taking ginkgo tablets helps your mental performance? Probability not.

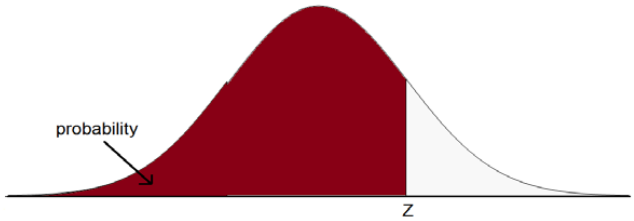
Remember that if we use $\alpha = .05$, we have a 5% probability of deciding the null hypothesis is false even if it is true. Or we will make a type I error 5% of the time.
So if we conduct 200 tests, we would expect to get $200\,(.05) = 10$ significant results even if the ginkgo tablets don't do anything.
So the fact that we got 9 significant results means it was probably those expected type I errors.

Standard Normal Table: Areas to the Left



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

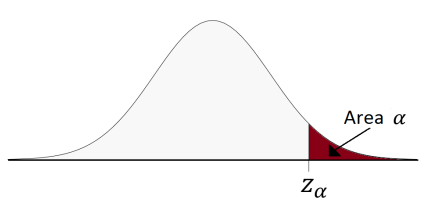
Standard Normal Table: Areas to the Left



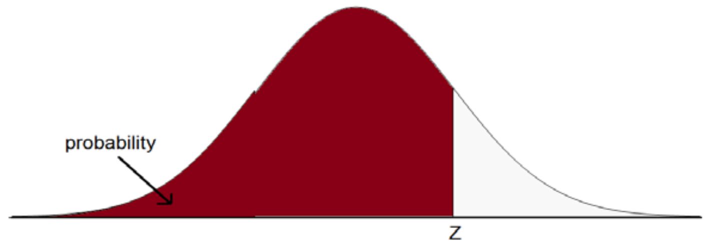
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.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
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Critical Points

α	z_α
0.10	1.282
0.05	1.645
0.025	1.960
0.01	2.326
0.005	2.576



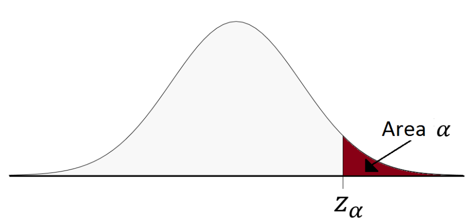
Standard Normal Table: Areas to the Left



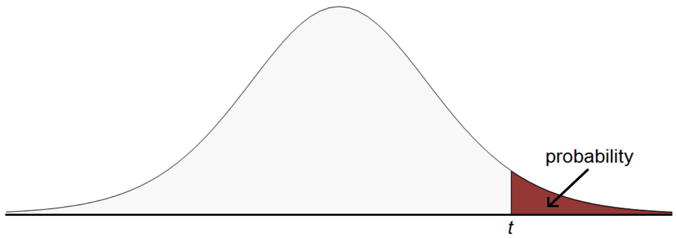
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.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
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Critical Points

α	z_α
0.10	1.282
0.05	1.645
0.025	1.960
0.01	2.326
0.005	2.576



T Table: Areas to the Right



degrees of freedom v	.20	.10	.05	.025	.01	.005	.001	.0005
	α							
1	1.376	3.078	6.314	12.71	31.82	63.66	318.3	636.6
2	1.061	1.886	2.920	4.303	6.965	9.925	22.33	31.60
3	0.978	1.638	2.353	3.182	4.541	5.841	10.21	12.92
4	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.862	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.861	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.860	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.859	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.858	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.858	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.857	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.856	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.856	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.855	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.855	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.854	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.854	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.851	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.849	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	0.848	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.846	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.845	1.290	1.660	1.984	2.364	2.626	3.174	3.390
200	0.843	1.286	1.653	1.972	2.345	2.601	3.131	3.340
∞	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291