Normal Distributions and Informal Inference

Previously, you

- Compared two or more distributions by assessing center, shape and spread
- Learned about methods of data collection and the importance of randomization
- Discovered the difference between observational studies and experiments and parameters and statistics

In this unit you will

- Use the Empirical Rule to find the percentage of data within one, two and three standard deviations of the mean
- Calculate z-scores to find the area underneath the curve for a particular score and given a percentile, find the score
- Create probability distribution tables
- Recognize a confidence interval and be able to interpret it
- Interpret data coming from surveys, observational studies and experiments to make decisions

You can use the skills in this unit to

- Fit a distribution with a normal model using technology
- Estimate the area under a normal curve and explain in context, with and without technology
- Determine whether experimental probabilities match given theoretical probabilities
- Determine if there is a significant difference between two treatments

Vocabulary

- **Bias –** A property of a statistical sample that makes sample results unrepresentative of the entire population.
- **Comparative Experiment** – They are designed to determine the differences between different forms of treatments.
- **Confidence Interval** – A range of values so defined that there is a specified probability that the value of a parameter lies within it.
- **Empirical Rule** – It states that for a normal distribution, 68% of the data will fall within one standard deviation of the mean, 95% of the data will fall within two standard deviations of the mean and 99.7% of the data will fall within three standard deviations of the mean.
- **Experimental (empirical) probability** – is the ratio of the number of times an event occurs to the total number of trials or times the activity is performed.
- **Hypothesis test** – When we use sample data to test whether there is convincing evidence for a claim about a population.
- **Margin of error** – An amount (usually small) that expresses the maximum expected difference between the true population parameter and a sample estimate of that parameter.
- **Normal distribution –** A distribution of values that is single-peaked (unimodal), symmetric and bell-shaped.
- **Parameter** – A measureable characteristic of a population.
- **Probability distribution –** A table that links each outcome of a statistical experiment with its probability of occurrence.
- **Sampling distribution –** A model that shows the results of taking many samples. It is usually in a table or in a histogram.

- **Simulation –** A model that uses the probabilities of a real-life situation.
- **Statistic –** A measurable characteristic of a sample.
- **Statistically significant –** The likelihood that a result or relationship is caused by something other than mere random chance.
- **Theoretical probability –** The number of ways that the event can occur, divided by the total number of outcomes. It is finding the probability of events that come from a sample space of known equally likely outcomes.

Essential Questions

- Why do we study normal distributions?
- Do experimental probabilities match the theoretical probabilities?
- How do you know when the difference between two treatments is statistically significant?
- There are many "studies out there", how do I know if they really are accurate?

Overall Big Ideas

In real life, data sets are large and many of these sets can be approximated by a normal distribution. Normal models allow us to answer and model real life situations.

Model probabilities found in an experimental environment and decide whether they are consistent with theoretical probabilities.

If a difference between the statistics of two treatments is outside of a critical confidence interval, the difference is statistically significant.

Understanding data by critically differentiating the merit of reports and data encountered in daily life.

Skills

To know and apply the empirical rule.

To use and apply normal distributions when appropriate (z-scores, %-tile).

To understand that inference is drawing a conclusion about a populatin parameter based on a random sample from that population.

To determine whether empirical results are consistent with the theoretical model.

To use data from a randomized experiment to compare two treatments.

To use simulations to decide if the difference between two statistics is significant.

To evaluate reports based on data.

Related Standards

S.ID.A.4
Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

S.IC.A.2
Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. For example, a model says a spinning coin falls heads up with probability 0.5. Would a result of 5 tails in a row cause you to question the model? *(Modeling Standard)

S.IC.B.4
Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling. *(Modeling Standard)

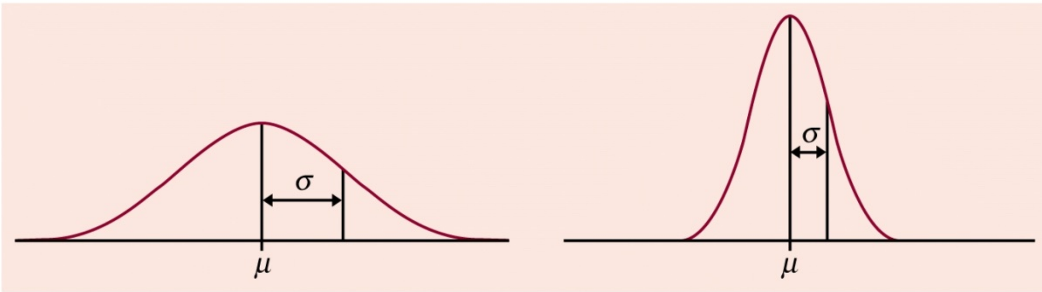
S.IC.B.5
Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant. *(Modeling Standard)

S.IC.B.6
Evaluate reports based on data. *(Modeling Standard)

Notes, Examples, and Exam Questions

<u>Sec. 9.4, 9.5 To know and apply the empirical rule and To use and apply normal distributions when appropriate (z-score, %-tile).</u>

Normal curves describe normal distributions. The curves are <u>single-peaked,</u> <u>symmetric</u> and <u>bell-shaped</u>. The curve is described by two things: its mean, μ, and its standard deviation, σ.

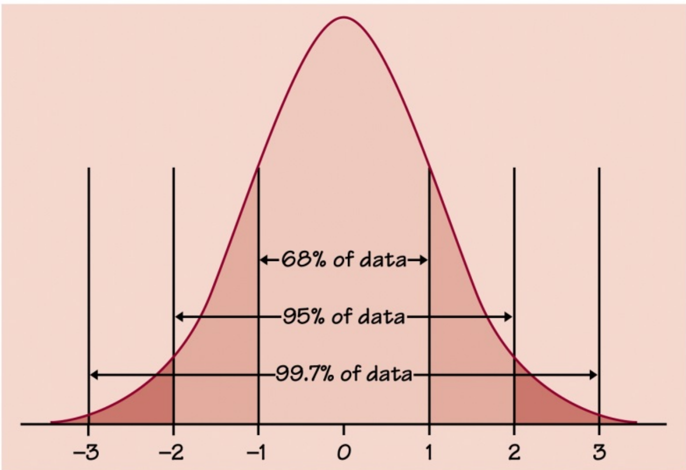


This is more spread out
This has a larger σ

This is less spread out
This has a smaller σ

Normal distributions are good descriptions of <u>real data</u>. Also, they are good approximations of many kinds of chance outcomes, like flipping a coin.

<u>The Empirical Rule (also known as the 68-95-99.7 Rule)</u>

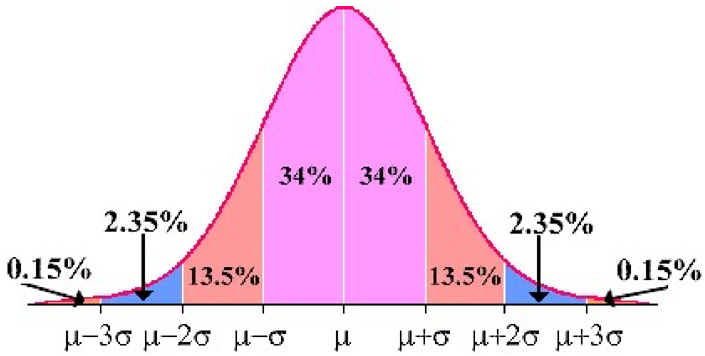


In a normal distribution, 68% of the observations fall within one σ of μ, 95% of the observations fall within two σ of μ, and 99.7% of the observations fall within three σ of μ.

Notation: $N(\mu,\sigma)$ If N(64.5,2.5), then we have a normal distribution with a mean of 64.5 and a standard deviation of 2.5.

Parameters are population variables. They are represented by Greek letters, like μ and σ. Statistics are sample variables. They are represented by Roman letters like $\overline{x}$ and s.

To break down the empirical rule further, divide into halves:



<u>Standardizing and z-scores:</u> If we have an observation, *x*, from a normal distribution with mean μ and standard deviation σ, the standardized value of *x* is: $z = \frac{x-\mu}{\sigma}$. A standardized value is called a **z-score**. *****A z-score tells us how many standard deviations the original observation falls away from the mean and in which direction.** These z-scores can be used to find the area under a normal curve. With the Empirical rule, one can only find the area for values that fall exactly one, two or three standard deviations from the mean.

Ex 1: Mike got an 85 on his science test. The class average was 72 and the standard deviation was 10. Mike also got an 89 on his math test and the math class's average was 80 with a standard deviation of 8. With respect to his courses, on which test did he perform better?

Find the z-scores for both tests: Science: $z = \frac{85-72}{10} = 1.30$ Math: $z = \frac{89-80}{8} = 1.13$

Interpretation: Mike scored 1.3 standard deviations above the science class mean and 1.13 standard deviations above the math class mean. Since, 1.3 > 1.13, Mike did relatively better on the science test.

The Standard Normal Distribution

The standard normal distribution is a normal distribution with a mean of 0 and a standard deviation of 1. $\mathrm{N}(\mu,\sigma) \to \mathrm{N}(0,1)$. All normal distributions can be standardized by using the z-score formula. This gives us a common scale.

Normal Distribution Calculations

Area under a density curve = proportion of the observations in the distribution. Once the distribution is standardized, find the areas using a normal probability table. The table is a table of areas under the standard normal curve. The table gives the area to the LEFT of the z score.

Using the z table

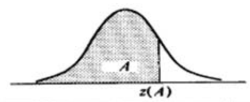
Find the z-score for the problem. Round to the nearest **hundredth**. Look up the z-score on the table. If you want the proportion to the left of that number, your answer is the four digit number you find. If you want the area to the right of that z-score, subtract the four-digit number from one (100%). Remember, the curve is symmetric, so if I want the area to the right of $z = 1.23$, that is the same as the area to the left of $z = -1.23$.

Continuous numbers

The normal distribution is describing continuous data. What this means is that the proportion of observations with $x > a$ is the same as the proportion with $x \geq a$. There is no area above a single point.

Part of the normal distribution table:

Entry is area A under the standard normal curve from $-\infty$ to $z(A)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.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

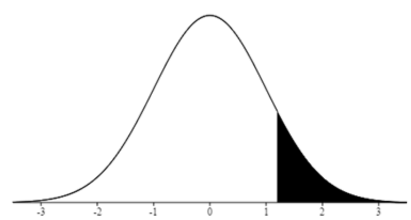
Ex 2: Sketch the curve and determine the following standard normal (z) curve areas .

a) The area under the *z* curve to the left of 1.75.

Look up 1.75 on the table: 0.9599

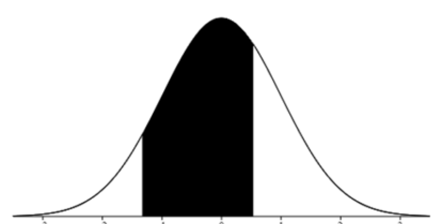
b) The area under the *z* curve to the right of 1.20.

Since the table only give us the values to the LEFT of the number, look up 1.20 and subtract from one: 1-0.8849 = 0.1151. OR look up -1.20 directly as the curve is symmetric and get: 0.1151.



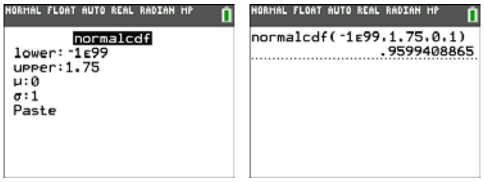
c) The area under the *z* curve between -1.33 and 0.53.

Since we are looking for the area between two values, look up both z-score values and subtract the two areas. 0.7019 – 0.0918 = 0.6101

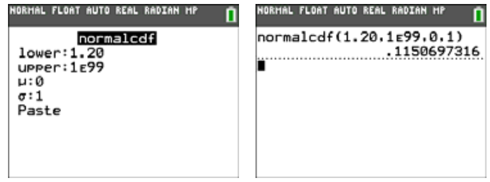


Using technology for example 2, use the normalcdf distribution on the TI-84 calculator. The input values are lower bound, upper bound, mean and standard deviation. Use 4 or more standard deviations away for lower and upper bounds at the edge of the curve. Press 2nd, VARS, 2 to find the normalcdf distribution.

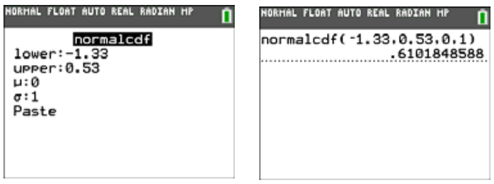
a)



b)



c)



Ex 3: Suppose the distribution of pH readings for soil samples taken in a certain geographical region can be approximated by a normal distribution with mean 6.00 and standard deviation 0.10. The pH of a randomly selected soil sample from this region is to be determined.

a) What is the probability that the resulting pH is between 5.90 and 6.00?

Use the Empirical Rule. The value 5.90 is exactly one standard deviation below the mean and according to the Empirical Rule, 34% of the values will be found in this area. Answer: 0.34

b) What is the probability that the resulting pH exceeds 6.20?

Using the Empirical Rule, 6.20 is exactly two standard deviations above the mean. 95% of the values fall within two standard deviations. Subtracting from one, we are left with 5% in the tails and dividing this value in half, that means 2.5% of the values fall in each tail. Answer: 0.025

c) What is the probability that the resulting pH is at most 5.95?

This requires z-scores as the Empirical Rule only works when the values are exactly one, two or three standard deviations away from the mean. Calculate the z-score and then use the table.

$$N(6,0.1): \quad P\left(x \le 5.95\right) = P\left(z \le \frac{5.95-6}{0.1} = -0.5\right) \to P\left(z \le -0.5\right) = \boxed{0.3085}$$

Ex 4: Let *x* denote the duration of a randomly selected pregnancy (the time elapsed between conception and birth). Accepted values for the mean value and standard deviation of *x* are 266 days and 16 days, respectively. Suppose that a normal distribution is an appropriate model for this distribution.

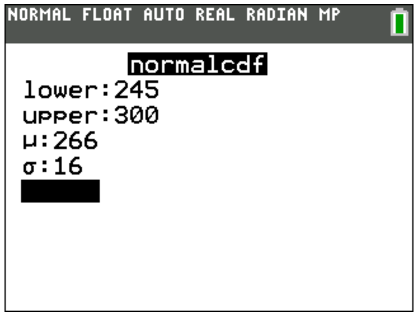
a) What percent of pregnancies last between 245 and 300 days?

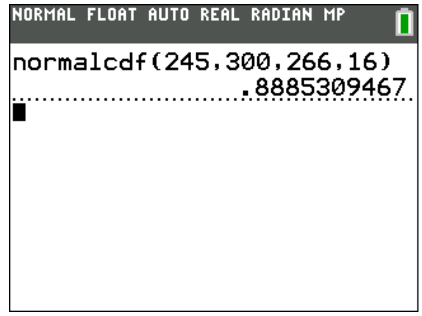
Calculate both z-scores:

$$N(266,16): \quad P\left(245 \le x \le 300\right) = P\left(\frac{245-266}{16} \le z \le \frac{300-266}{16}\right)$$

$$= P\left(-1.31 \le z \le 2.13\right) = 0.9834 - 0.0951 = \boxed{0.8883}$$

Calculator:





b) What percent of pregnancies have a duration less than 230 days?

$$P(x \le 230) = P\left(z \le \frac{230-266}{16} = -2.25\right) \to P(z \le -2.25) = \boxed{0.0122}$$

c) What percent are within 16 days of the mean duration?

Since 16 days is the standard deviation, we can use the Empirical Rule. Find the percent of values that would be within one standard deviation of the mean: 68%.

d) A *Dear Abby* column dated Jan. 20, 1973 contained a letter from a woman who stated that the duration of her pregnancy was exactly 310 days. (She wrote that the last visit with her husband, who was in the Navy, occurred 310 days prior to birth.) What percent of pregnancies are at least 310 days? Does this make you a bit skeptical of the claim?

$$P(x > 310) = P\left(z > \frac{310-266}{16} = 2.75\right) \to P(z > 2.75) = \boxed{0.0030}$$. I would be very skeptical of this claim as this is a very small percentage.

Finding the value when given the area:

It is possible to find the z-score if we are **given the proportion** or area by working backwards. Look in the MIDDLE of the table for the CLOSEST decimal related to the percent given and then read OUT to find the z-score. You are are reading the table "backwards" from what was being done in examples 2, 3 and 4. Example: Because $P(z < 0.44) = 0.67$, sixty-seven percent of all z values are less than 0.44, and 0.44 is the 67^{th} percentile of the standard normal distribution. Once the z-score is found from the table, use the formula to solve for *x*. To transform the z-score formula solved for *x*, we get: $x = z\sigma + \mu$.

Ex 5: Determine the value of each of the following percentiles for the standard normal distribution.

a) 75th percentile

Look up the closest four-digit number to 0.7500 on the table. Read out to find the z-score. The closest value to 0.7500 is 0.7486. Reading up and left, the z-score is 0.67. If you are at the 75th percentile, that means you are about 0.67 standard deviations above the mean.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	[illegible]	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	[illegible]	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.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	[illegible]	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	[illegible]	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319

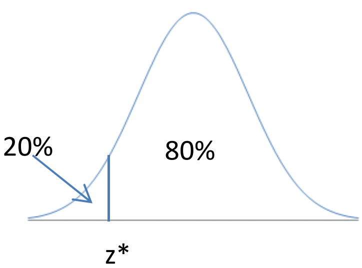
b) 90th percentile

Using the table above, the closes number to 0.9000 is 0.8997, giving us a z-score of 1.28.

c) 40th percentile

Since this is below the 50th percentile (the mean), look at the negative side of the normal distribution table. Find the closest four-digit number to 0.4000 which is 0.4012 or a z-score of -0.25.

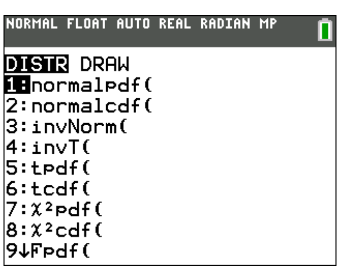
Ex 6: The light bulbs used to provide exterior lighting for a large office building have an average lifetime of 700 hours. If the distribution of the variable x = length of bulb life can be modeled as a normal distribution with a standard deviation of 50 hours, how often should all of the bulbs be replaced so that only 20% of the bulbs will have already burned out?

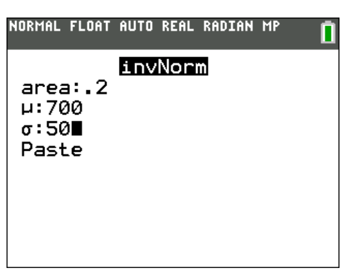


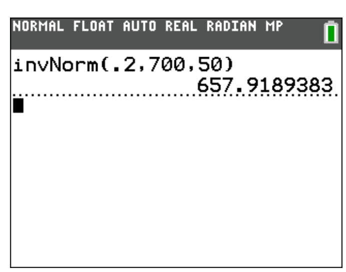
The closest four-digit value to 0.200 is 0.2005 with a z-score of -0.84.

$$N(700,50):\ x = z\sigma + \mu \rightarrow x = (-0.84)(50) + 700 = \boxed{658 \text{ hours}}$$

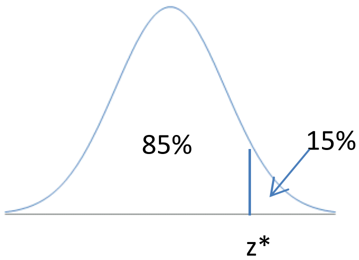
Technology can also be used to solve the problem. With the TI-84 calculator, the invNorm command can be used. To get to the invNorm command, press 2nd, VARS, 3. The required parameters for the command are mean, standard deviation and area to the LEFT of the value.







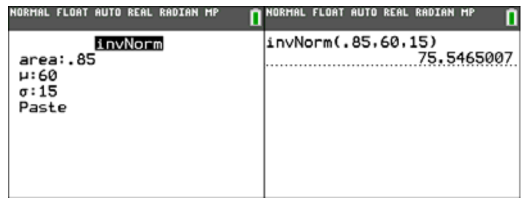
Ex 7: **Your Statistics teacher has just given an exam. The grade distribution of the exam can be approximated by the normal distribution with a mean of 60 and a standard deviation of 15. If the top 15% of the tests will receive an A, what scores will qualify?**



The closest four-digit value to 0.8500 is 0.8508 with a z-score of 1.04.

$N(60,15):\ x = z\sigma + \mu \rightarrow x = (1.04)(15) + 60 = \boxed{75.6}$

Any student having a score of 76 or higher will receive an A.



Ex 8: **Consider the variable x = time required for a college student to complete a standardized exam. Suppose that for the population of students at a particular university, the distribution of x is well approximated by a normal curve with mean 45 min and standard deviation 5 min.**

a) If 52 minutes is allowed for the exam, what percent of the students would be unable to finish in the allotted time?

$$N(45,5):\ P(x>52) = P\left(z > \frac{52-45}{5} = 1.40\right) \rightarrow P(z>1.40) = \boxed{0.0808} \rightarrow \text{ about 8\%.}$$

b) How much time is required for the fastest 25% of all students to complete the exam?

Look up 0.7500 for the top 25% and the closest value is 0.7486 which gives us a z-score of 0.67.

$$N(45,5):\ x = z\sigma + \mu \rightarrow x = (0.67)(5) + 45 = \boxed{48.35 \text{ minutes}}.$$

SAMPLE EXAM QUESTIONS

1. The mean is 80 and the standard deviation is 10. What is the standard score of an observation of 90?

A. 90

B. 0

C. 10

D. 1.0

E. -1.0

Ans: D

2. A normal distribution always

A. Is skewed to the right

B. Is skewed to the left

C. Is symmetric

D. Has a mean of 0

E. Has more than one peak

Ans: C

The next four questions use this information: The length of pregnancy isn't always the same. In pigs, the length of pregnancies varies according to a normal distribution with mean 114 days and standard deviation 5 days.

3. What range covers the middle 95% of pig pregnancies?

A. 109 to 119

B. 104 to 124

C. 99 to 129

D. 94 to 134

Ans: B

4. What percent of pig pregnancies are longer than 114 days?

A. 16%

B. 34%

C. 50%

D. 84%

Ans: C

5. What percent of pig pregnancies are longer than 109 days?

A. 16%

B. 34%

C. 50%

D. 84%

Ans: D

6. The median length of a pig pregnancy is

A. 119 days

B. 114 days

C. 109 days

D. Between 109 and 119 days, but cannot be more specific

E. Greater than 114 days, but cannot be more specific

Ans: B

7. Scores on the American College Testing (ACT) college entrance exam follow the normal distribution with mean 18 and standard deviation 6. Jon's standard score on the ACT was -0.7. What was Jon's actual ACT score?

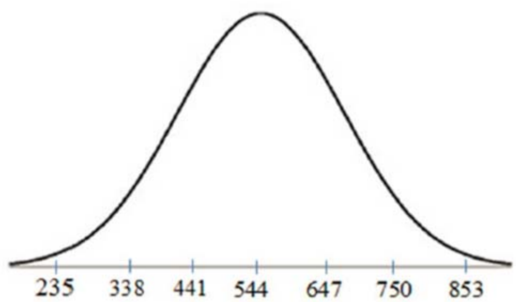
A. 4.2 **C.** -4.2 **E.** 13.8

B. 22.2 **D.** 9.6

Ans: E

8. The Graduate Record Examinations are widely used to help predict the performance of applicants to graduate schools. The range of possible scores on a GRE is 200 to 900. The psychology department at a university finds that the scores of its applicants on the quantitative GRE are approximately Normal with mean = 544 and standard deviation = 103.

(1) Make an accurate sketch of the distribution of these applicants' GRE scores. Be sure to provide a scale on a horizontal axis.



(2) Use the 68-95-99.7 rule to find the proportion of applicants whose score is between 338 and 853.

853 is 3 standard deviations above 544, so 99.7/2 or 49.85% of the scores are between 544 and 853. 338 is 2 standard deviations below the 544, so 95/2 or 47.5% of the scores are between 338 and 544. Thus, 49.85 + 47.5 = 97.35% of the scores are between 338 and 853.

(3) What proportion of GRE scores are below 500?

$$N(544,103): \;\; P(x<500)=P\left(z<\frac{500-544}{103}=-0.43\right)$$

$z<-0.43=0.3336$ or 33.36% of the scores below 500

(4) What proportion of GRE scores are above 800?

$N(544,103):\ \ P(x>800)=P\left(z>\dfrac{800-544}{103}=2.49\right)$

$z>2.49=0.0064$ or 0.64% of the scores above 800

(5) Calculate and interpret the 34th percentile of the distribution of applicants' GRE scores.

The 34th percentile corresponds to $z=-0.41$

$x=(-0.41)(103)+544$

$x=502$ So about 34% of the applicants have GRE scores below 502.

9. There are two major tests of readiness for college, the ACT and the SAT. ACT scores are reported on a scale from 1 to 36. The distribution of ACT scores in recent years has been roughly Normal with mean μ=20.9 and standard deviation σ=4.8. SAT scores (prior to 2005) were reported on a scale from 400 to 1600. SAT scores have been roughly Normal with mean μ=1026 and standard deviation σ=209. The following exercises are based on this information.

(1) Joe scores 1245 on the SAT. Assuming that both tests measure the same thing, what score on the ACT is equivalent to Jose's SAT score? Explain.

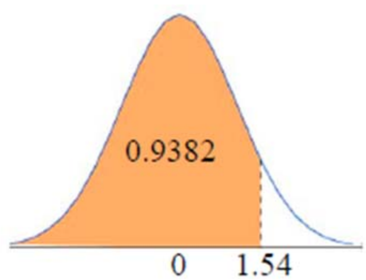
$z=\dfrac{1245-1026}{209}\rightarrow z=1.05$

$x=(1.05)(4.8)+20.9=25.94$ The equivalent score on the ACT is 25.94.

(2) Reports on a student's ACT or SAT usually give the percentile as well as the actual score. Terry scores 1342 on the SAT. What is her percentile? Show your method.

$z=\dfrac{1342-1026}{209}=1.54$

Terry is at the 94th percentile.



(3) The quartiles of any distribution are the values with cumulative proportions 0.25 to 0.75. What are the quartiles of the distribution of ACT scores? Show your method.

25th percentile

$z=-0.67$

$-0.67=\dfrac{x-20.9}{4.8}$

$x=17.7$

75th percentile

$z=0.67$

$0.67=\dfrac{x-20.9}{4.8}$

$x=24.1$

Units 9.7 – 9.12

To understand that inference is drawing a conclusion about a population parameter based on a random sample from that population.
To determine whether empirical results are consistent with the theoretical model.
To use data from a randomized experiment to compare two treatments.
To use simulations to decide if the difference between two statistics is significant.
To evaluate reports based on data.

These units are meant to be an INFORMAL introduction to inference. The students are not expected to construct confidence intervals or do tests of significance. This unit is to prepare students for these topics which they will see in AP Statistics. Concentrate on familiarizing students with the vocabulary, have students do simulations to construct models, and look at real-world reports on data and discuss and evaluate the findings.

What is Statistics? Statistics is the science of conducting studies to collect, organize, summarize, analyze and draw conclusions from data. The two main branches of statistics are descriptive statistics and inferential statistics. Descriptive statistics consists of the collection, organization, summarization and presentation of data. Inferential statistics consists of the generalizing from samples to populations, performing estimations and hypothesis tests, determining relationships among variables, and making predictions. Statistical inference provides methods for drawing conclusions about a population from sample data. We use probability to express the strength of our conclusions. The two most common types of formal statistical inference are confidence intervals and tests of significance. Both report probabilities that state what would happen if we used the inference method many times.

Vocabulary:

1) **Parameter –** A numerical characteristic of a population. The population can be of any size but encompasses all the people/objects that we want information about. Greek letters are usually used to represent a parameter. Examples: Mean - $\mu = \text{the true mean of the population}$; Proportion - $p = \text{the true proportion of the population}$. Some texts will use π for p to represent the population proportion.

2) **Statistic -** A numerical characteristic of a sample. A sample is a set of items that are drawn from the population. Since the only reason to ever draw a random sample is to infer something about the population from which it came, it should be clear that when a statistic is calculated it is done in order to estimate a corresponding parameter of the population from which the same was drawn. Roman letters are usually used to represent a statistic. Examples: Mean - $\overline{x} = \text{the mean of the sample}$; Proportion - $\hat{p} = \text{the proportion of the sample}$. If the text uses, π, then it will use p for the sample proportion instead of $\hat{p}$ (p-hat).

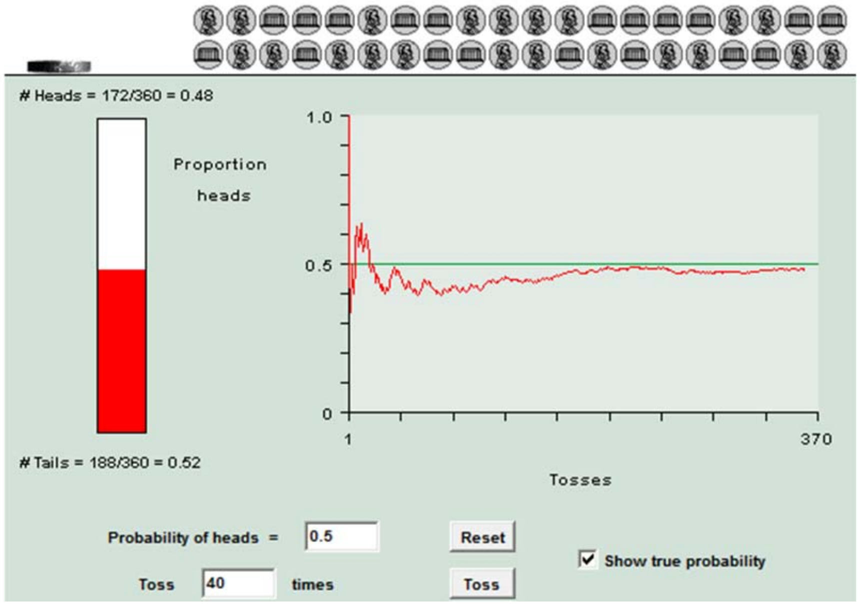
3) **Unbiased –** A statistic is said to be an unbiased estimate of a given parameter when the statistic can be shown to be equal to the parameter being estimated. If information for the sample is gathered randomly and avoiding other biases, we say that $\overline{x} = \mu$.

4) **Random –** What is randomness? Randomness means lack of pattern or predictability in events. Some things that are random are rolling dice, shuffling cards, flipping a coin and bingo. We use randomness in our data collection to give a fair and accurate picture of the population (unbiased). Drawing conclusions from data relies on randomness in data collection. Inferential

procedures cannot be performed unless the sampling was done randomly or the experiment contained random assignment.

5) **Theoretical vs. Experimental Probability -** Theoretical probability is the mathematically calculated value. For example, for flipping a coin three times, the probability of getting three tails is $(0.5)^3 = 0.125$. Experimental or Empirical probability is the outcome of a simulation or an actual physical experiment. An example would be to flip a coin three times and record how many tails come up. Do this ten times. Out of the ten times, one trial had all three tails, so the experimental probability is $\frac{1}{10} = 0.10$. In this case, the experimental and theoretical probabilities are very similar.

6) **Law of Large Numbers –** States that the more times a random experiment is repeated, the closer the average of that experiment will be to the expected value. This is important because it "guarantees" stable long-term results for the average of some random events. Example: If we flip a coin ten times, there is only a 0.246 probability of getting 5 heads and 5 tails. It is possible we toss and get 2 heads and 8 tails. However, if we flip a coin ten thousand times, it will be more likely that we see approximately 5000 heads and 5000 tails.



http://bcs.whfreeman.com/ips4e/cat_010/applets/Probability.html

Simulations: If you toss a coin ten times, what is the probability of having a run of three or more "heads" in a row? If an airline "overbooks" a certain flight, what is the chance more passengers show up than the airplane has seats for? When 67 people get cancer in the 250 homes in a small town, could that be chance alone, or is polluted well water a more likely explanation of the cluster of cancer cases? When the mathematics becomes too complex to figure out the theoretical probability of certain events, statisticians often use simulations instead. Simulations can also be used to check statistical computations, or they can be used in place of a study that is too expensive, time-consuming, or unethical. A simulation is a model that uses the probabilities of a real-life situation.

A **simulation** consists of a collection of things that happen at random. The most basic event is called a **component** of the simulation. Each component has a set of possible outcomes, one of which will occur at random. The sequence of events we want to investigate is called a **trial**. Trials usually involve several components. After the trial, we record what happened, which is our **response variable**. There are seven steps to a simulation.

1. Identify the component to be repeated.
2. Explain how you will model the outcome.
3. Explain how you will simulate the trial.
4. State clearly what the response variable is.
5. Run several trials (as many as possible).
6. Analyze the response variable.
7. State your conclusion (in the context of the problem).

Ex 9: Simulation example: Suppose that we have a basketball player who is an 80% free-throw shooter. How many shots can she make in a row without missing?

- **Component:** the most basic event we are simulating – a single free-throw (foul shot)
- **Outcome:** The player makes 80% of her free-throws, so to simulate 80% using a random number table, let the digits 0 – 7 represent a shot that is made and 8 and 9 will represent missed shots.
- **Trial:** the sequence of events we want to investigate: shooting until a miss. Using the random number table, read digits until a failure (8 or 9) occurs. Then, count the number of success until a failure occurred. This is the end of one trial. Record that number and repeat.
- **Response variable:** The number of successes (made shots)
- **Statistic:** Find the mean number of shots made. If 50 trials were run, find the average number of successes for the 50 trials.
- **Conclusion:** On average, according to our simulation, she is expected to make _____ shots in a row without missing.

Ex 10: Identify the values as a parameter or a statistic and use proper variable notation.

According to Snapple.com, **13%** of adults are left-handed. At a school administrator's conference, **16%** of those attending were left-handed.

13% is the parameter $p = 0.13$

16% is the statistic $\hat{p} = 0.16$

Ex 11: Mr. and Mrs. Smith want to have children and would love to have a girl, but they don't want to have more than four children. They want to figure out the chances of having a girl if they have children until they have a girl or until they have four children, whichever comes first.

a) Set up the simulation

Component – Simulating having a girl.

Outcome – Having a girl has a 50% chance, so use a coin to simulate. Heads = girl, Tails = boy

Trial – Flip the coin until you get "heads" or until you have flipped four times. Record whether a girl was born.

Response variable – The trials that contain a girl birth

Statistic – Find the proportion of the 25 trials that contain a girl birth.

b) Run the simulation 25 times and tally the results

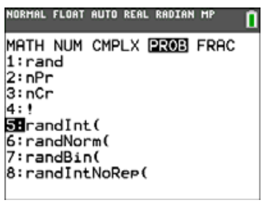
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y

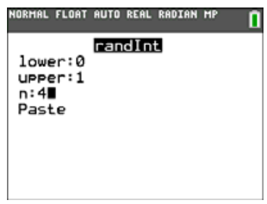
c) Combine results with the rest of the class and answer: What is the couple's probability of having a girl in this situation?

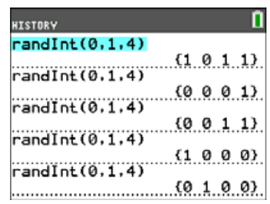
For the situation above, 23/25 or 92% of the trials had a girl birth. Answers will vary as all students' results are combined.

d) Repeat the simulation using the random number generator on the calculator.

Let "0" represent a boy and "1" represent a girl. Use the random integer generator on the TI-84 calculator where the first argument is the starting number, the second argument is the ending number and the third argument is the number of random numbers to generate. Keystrokes: MATH, PRB, 5. Hit enter to get a new set of numbers. Note that in the five trials shown below, a girl (a "1") was born in each trial, so all 5 were successes.







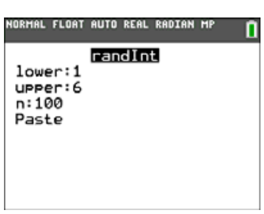
e) Calculate the theoretical probability for the situation above.

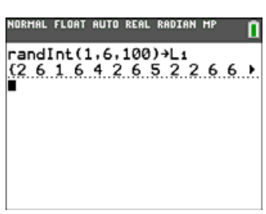
Girl on the first birth: P(G) = 0.5
Girl on the second birth: P(B and G) = (0.5)(0.5) = 0.25
Girl on the third birth: P(B and B and G) = (0.5)(0.5)(0.5) = 0.125
Girl on the fourth birth: P(B and B and B and G) = (0.5)(0.5)(0.5)(0.5) = 0.0.0625
Sum = 0.5+0.25+0.125+0.0625 = <u>0.9375</u>. Note that this is close to the experimental probability found in c).

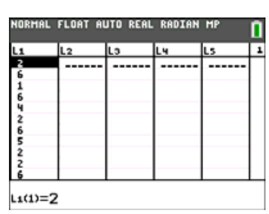
Ex 12: Katelyn is going to be babysitting her nephew a lot this summer. She has the great idea that one way to entertain him is to walk to McBurger's for a Kids' Meal for lunch. The Kids' Meal comes packed randomly with one of six equally-likely action figures. Katelyn is worried that her nephew will be disappointed unless he gets all six figures. She wants to know how many meals she can expect to buy before he will get all six figures.

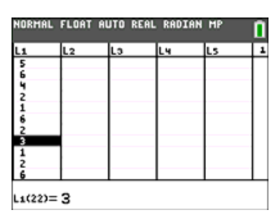
a) How can this situation be modeled with a simulation?

This can be simulated using one die and rolling until all six numbers come up. Record how many rolls were needed to acquire all six numbers. This can also be done using a random number table. Read single digits at a time (ignoring 7, 8, 9 and 0) and count how many digits it takes until 1 – 6 have all come up. Using a random number generator, generate numbers from 1 through 6 and generate a large amount and count how many it takes to acquire all six numbers. Store the numbers in a list to make reading easier. Below it took 22 numbers before all six numbers were generated.









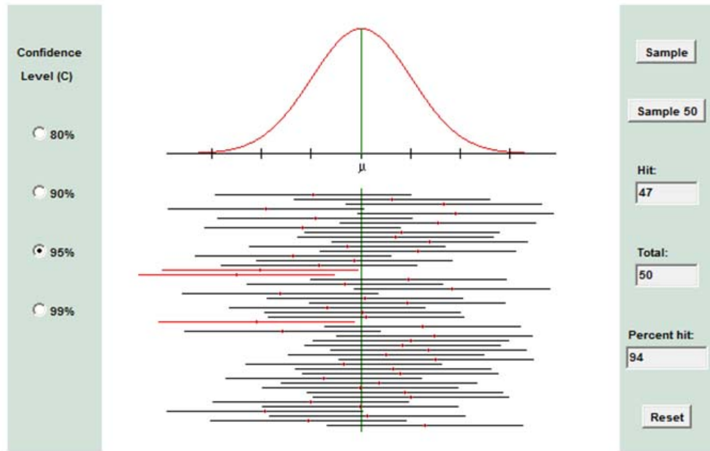
b) Run the simulation ten more times. Keep track of how many Kid's Meals it takes to get all six action figures. Calculate your *wait time*, or the average number of meals it took to get all six figures. Is your value close to the theoretical probability?
Answers will vary. In the long run, it should be close to the theoretical probability which is 14.7.

c) Estimate a range for the number of meals you might need to buy. What is the most you need to buy? The least? Do you think it possible that you may have to buy 50 meals to get all six action figures? 100 meals?
Buying 50 or 100 meals before getting all six action figures is possible, but not very probable.

Margin of error: If a sample is selected randomly, taking care to avoid bias, we can be confident that the sample represents the whole population. Making inferences from samples about whole populations is at the heart of statistics. Let's say your friend Maria says that you have to have the latest, hottest iPhone – everybody else already has it! What proportion of young people have the iPhone? We cannot possibly ask everybody in the United States who is between the ages of 12 and 20, but we can take a random sample of this population and calculate the proportion that have the phone. If in a random sample of 1000 members of our population, 250 have the iPhone, it still does not mean that exactly ¼ or 25% of all people between the ages of 12 and 20 in the U.S. have the iPhone. The proportion of young people with the phone will naturally vary from sample to sample – some samples of young people will have more, some will have less phones. This is called **sampling variability** and it is unavoidable. If we knew how much the proportion naturally varied from sample-to-sample, we could establish a range of estimates for the proportion of phones in the teen population. This **margin of error** is frequently reported in statistical studies. You might read in a newspaper that the percentage of teens with the current iPhone is 23% with a margin of error of 3%. There will also be a probability associated with this number and it is usually 95%. This means that statisticians can be 95% confident that based on their small sample, that between 20% and 26% of all teens own this new iPhone. Note that the **margin of error** is half the spread between the upper and lower bounds.

What margin of error means: "Margin of error plus or minus 3 percentage points" is shorthand for this statement: If we took many samples using the same method we used to get this one sample, 95% of the samples would give a result within plus or minus 3 percentage points of the truth about the population. What does this mean? A sample chosen at random will usually not estimate the truth about the population exactly. We need a margin of error to tell us how close our estimate comes to the truth. But we can't be *certain* that the truth differs from the estimate by no more than the margin of error. Ninety-five percent of all samples come this close to the truth, but 5% miss by more than the margin of error. We don't know the truth about the population, so we don't know if our sample is one of the 95% that capture it or one of the 5% that miss. We say we are **95% confident** that the truth lies within the margin of error. Why 95%? The more confident we want to be (like 99%), the wider the interval will have to be and we lose precision. If we lower the confidence (like 90%0, the interval will be narrower, but we lose accuracy. 95% is a good compromise between the two.

In the applet, fifty confidence intervals were constructed. The dot represents the sample statistic, which is the center of the interval. The margin of error is added on to each side of the statistic. The lines on each side of the dot span the confidence interval. Note that in these 50 samples, 47 captured the true parameter and 3 did not – that is a 94% percent hit.



http://bcs.whfreeman.com/ips4e/cat_010/applets/confidenceinterval.html

Ex 13: A *New York Times* poll on women's issues interviewed 1025 women randomly selected from the United States. One question was "Many women have better jobs and more opportunities than they did 20 years ago. Do you think women had to give up too much in the process, or not?" Forty-eight percent of women in the sample said, "yes".

a) The poll announced a margin of error of $\pm 3\%$ for 95% confidence in its conclusions. Make a confidence statement for this situation.

Based on this sample, I am 95% confident that between 45% and 51% of adult women feel that women had to give up too much.

b) Explain to someone who knows no statistics why we can't just say that 48% of all adult women feel that women had to give up too much.

The statistic 48% came from one sample and because of sampling variability, another unbiased random sample of that same population will probably give us a different result. It is better to report a range of values with a specified confidence level to better estimate the population parameter and to account for the sample-to-sample variability.

c) Explain clearly what "95% confidence" means.

It means that if repeated samples were taken and the 95% confidence interval computed for each sample, 95% of the intervals in the long run would contain the true population parameter. Naturally, 5% of the intervals would not contain the population parameter. (See figure above)

Ex 14: The margin of error for a poll is 4%. This means that

(a) we have 95% confidence that the sample statistic is within 4% of the population parameter
(b) 4% of those sampled did not answer the question asked
(c) 4% of those sampled gave the wrong answer to the question asked
(d) 4% of the population were in the sample
(e) the confidence we have in the statistic is 4%

Ans: A

Ex 15: A poll of 1190 adults finds that 702 prefer balancing the budget over cutting taxes. The announced margin of error for this result is plus or minus 4 percentage points. The news report states the confidence level is 95%.

a) What is the value of the sample proportion $\hat{p}$ who prefer balancing the budget? Explain in words what the population parameter *p* is in this setting.

$\hat{p} = \frac{702}{1190} = 0.59 = 59\%$. ***p* is the proportion of all adults who prefer balancing the budget over cutting taxes.**

b) Make a confidence statement about the parameter *p*.

I am 95% confident that between 55% and 63% of all adults prefer balancing the budget over cutting taxes. Since all values in the confidence interval are above 50%, I am confident that a majority prefer balancing the budget.

Ex 16: Sample-to-sample variability to determine whether two results in an experiment are truly different.

Endangered Frogs: The female red-eyed poison-dart frog visits numerous bromeliads ("air plants") where pools of water have collected in the leaves. In each pool, she lays a single egg that grows into a tadpole. The tadpoles feed on mosquito larvae in the pools, but local health officials have been killing the mosquito larvae. Environmentalists are concerned because they are unsure whether the tadpoles will adapt to eat some other food source.

Luis led a team of environmental scientists to the Costa Rican rain forest. They tagged 100 female tadpoles and determined the number that grew to adulthood. Mosquito larvae were placed in 50 bromeliads, and the other 50 bromeliads were treated so that mosquitoes could not live there.

36 of the 50 tadpoles (72%) with mosquitoes survived, while 29 of the 50 (58%) non-mosquito tadpoles survived. The difference in the proportion that survived is 0.72 – 0.58 = 0.14. Investigate whether a difference in the proportion of tadpoles that survived of 0.14 can be explained by natural sample-to-sample variability, or if there is a true difference between the two groups.

First, explore the sample-to-sample variability using a computer/calculator simulation.

a) 72% of the mosquito-fed tadpoles survived. For the simulation, the numbers 1 to 72 can represent the mosquito-fed tadpoles that survived. The numbers 73 to 100 will represent a tadpole that died. Simulate one sample of 50 mosquito-fed tadpoles. What proportion of your sample survived?

Answers will vary around 0.72

b) 58% of the non-mosquito tadpoles survived. Simulate one sample of 50 non-mosquito tadpoles. What proportion of your sample survived?

Answers will vary around 0.58

c) What is the *difference* between these two proportions (mosquito-fed proportion that survived minus non-mosquito proportion survived)?

Answers will vary around 0.14, but could be negative.

d) The difference of proportions in your random sample that survived is likely to be higher or lower than 0.14, or even negative, because of natural sample-to-sample variability. What does a positive difference mean? What would a negative difference mean?

A positive difference means that more survived than did not; a negative difference means more died than survived.

e) If we take repeated random samples by repeating the simulation, we can estimate the sample-to-sample variability, and then the margin of error. Below are the results of 100 trials. The table shows the differences in the proportion that survived (mosquito-fed proportion that survived minus non-mosquito proportion survived).

Proportion survived minus proportion did not survive. Simulation mean = 0.14				
0.10	0.33	0.13	0.13	0.10
0.05	0.10	0.02	–0.05	0.14
0.21	0.12	0.22	0.22	0.17
0.13	0.24	–0.07	0.23	0.14
0.16	0.30	0.12	0.21	0.11
0.17	0.19	0.14	0.13	0.09
0.05	0.10	0.23	0.29	0.30
0.26	0.08	0.07	0.30	0.05
0.12	–0.04	0.16	0.03	0.13
0.26	0.16	0.08	0.14	0.07
0.27	0.19	–0.01	0.19	0.16
0.06	0.17	0.18	0.06	0.13
0.15	0.11	0.07	0.05	0.15
0.18	0.24	–0.02	0.10	0.16
0.19	0.18	0.36	0.20	0.06
0.22	0.17	0.16	0.26	–0.04
0.13	0.05	0.05	0.15	0.19
0.13	0.18	0.10	0.13	0.10
0.16	0.21	0.17	0.09	0.03
0.13	0.13	0.12	0.15	0.23

Find an upper and lower bound on the middle 90% of the proportions and the margin of error.

The lower bound is –0.015, the upper bound is 0.295, and the margin of error is (0.295 + 0.015 = 0.31) and 0.31/2 = ± 0.155.

f) You have found the sample-to-sample variability. You are ready to make a prediction for the entire population. Report your prediction for the difference between the proportion of mosquito-fed tadpoles that survived and the proportion of non-mosquito tadpoles that survived.

0.14 ± 0.155 = (-0.02, 0.30). That means, we are 90% confident that the true proportion of mosquito-fed tadpoles that survived is anywhere from 2% less than the non-mosquito tadpoles to 30% more.

g) What does a difference of zero mean in the context of this problem?

A difference of zero means that there is no difference in the survival rate of tadpoles that ate mosquitoes and tadpoles that did not.

h) Are you convinced (do you have evidence) there is a true difference in the tadpoles that ate mosquitoes and those that did not? Explain why or why not.

No. A difference of zero is included within the margin of error. A difference of zero is a plausible result for the population of all tadpoles. We are not convinced there is a true difference in survival between tadpoles that ate mosquitoes and those that did not.

Hypothesis Tests: In Example 16, the Endangered Frogs problem, we informally looked at a hypothesis test. We had two samples and we calculated the difference between those two samples. Then by looking at the sample-to-sample variability of the differences between the two populations, a margin of error was created for the true difference between the two populations. If a difference of zero was within that margin of error, it can be concluded that it was plausible that there was no difference between the two populations; it cannot be concluded that the two populations were in fact different. On the other hand, if a difference of zero is not within the margin of error, it can be concluded that it was plausible that there was actually a true difference between the two populations.

Significance or hypothesis testing is used to help make a judgment about a claim by addressing the question, "Can the observed difference be attributed to chance? A statistical hypothesis is an assumption about a population parameter. This assumption may or may not be true. Hypothesis testing refers to the formal procedures used by statisticians to accept or reject statistical hypotheses. The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.

The significance test can be broken down into four steps:

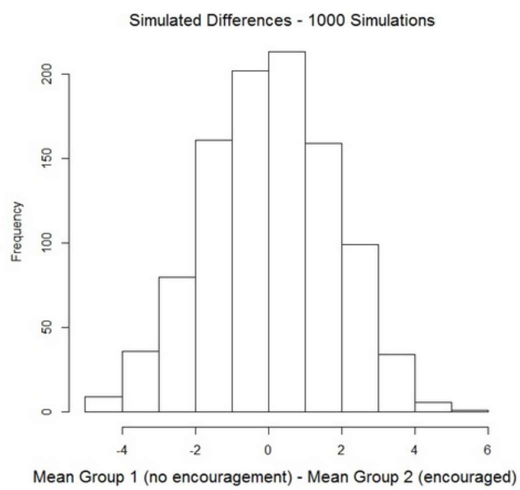
1. **Null and alternative hypotheses** – The null hypothesis, H_0, is a claim of "no difference". It is the statement of no change. The opposing hypothesis is the alternative hypothesis, H_a, is a claim of "a difference in the population" and is the hypothesis the researcher often hopes to bolster.
2. Formulate an analysis plan. The analysis plan describes how to use sample data to evaluate the null hypothesis. The evaluation often focuses around a single test statistic.
3. Analyze the sample data. Find the value of the test statistic. Large test statistics indicate data are far from expected, providing evidence against the null hypothesis and in favor of the alternative hypothesis.
4. Interpret the results. The test statistic is converted to a conditional probability called a ***p*-value**. The *p*-value answers the question, "If the null hypothesis were true, what is the probability of observing the current data or data that is more extreme?" **Small *p* values** provide evidence against the null hypothesis (reject the null) because they say the observed data are unlikely when the null hypothesis is true. The use of "significant" in this context means "the observed difference is not likely due to chance". It does NOT mean of "important" or "meaningful". What is **small**? Most times we compare the *p*-value to 0.05. If the *p*-value is smaller than 5%, we reject the null hypothesis and say that the data is statistically significant and if the *p*-value is larger, we fail to reject the null hypothesis and say the results are not significant.

Example of a Test: Suppose we want to determine whether a coin was fair and balanced. A null hypothesis might be that half the flips would result in Heads and half, in Tails. The alternative hypothesis might be that the number of Heads and Tails would be very different. Symbolically, this would be expressed as: $H_0: p = 0.5$ and $H_a: p \neq 0.5$. The analysis plan would be to do the simulation. Flip the coin 50 times and let's say that the result was 40 Heads and 10 Tails. The proportion of heads is 0.80. Given this result, we would be inclined to reject the null hypothesis because what was expected (0.50) and what was observed (0.80) are not close in value. We would then conclude, based on the evidence that there is significant statistical evidence to reject the null hypothesis and say that the coin was not fair and balanced.

<u>How deep should you go?</u>

At this level, we are not expecting the Algebra II students to apply hypothesis testing or confidence interval calculations formally. Instead, we want them to start thinking about the meaning behind these procedures before they see them formally presented at the college level or in an AP Stats course. These types of ideas will help the students have a much better idea of the p-value and the whole process of hypothesis testing itself, once these are introduced. Here is an example of what students should be able to do.

Ex 17: Suppose that two researchers want to determine if high school students that are offered encouraging remarks complete a difficult task faster, on average, than those who aren't. In order to test this, they select two random samples of 25 high school students each. The first group is asked to work on a difficult puzzle and offered no feedback as they work. The second group is asked to do the same but are also given encouraging comments such as "you almost got it" or "that's a good idea" as they work. For the first group (no encouragement), the mean time to complete the puzzle was 28.1 minutes with a standard deviation of 6.7 minutes. For the second group, the mean time was 27.2 minutes with a standard deviation of 5.5 minutes. In order to test the significance of this result, the researchers used a computer to randomly assign individual times to each group and then compute the new mean difference between the first and second groups. They then repeated this process 1,000 times and plotted all of the resulting differences on the plot below.



The question here might be then to ask students to determine if the observed difference between the means is statistically significant, or explain whether or not this should lead researchers to believe that those with encouragement will complete the task faster. These are deep/critical thinking types of questions that go beyond applying a formula.

Using the graph, we would hope that they would see that the observed difference of 28.1 – 27.2 = 0.9 minutes is within a range of values that is frequently observed when the groups are assigned randomly (it is not a rare difference – it came up a lot in simulation). Therefore, the experiment's results are not statistically significant as they could be due to chance alone. Through resampling, they are able to see how the samples might behave if the differences WERE due to chance (as they were in the simulation).

As you can see, this type of question is indirectly having students think about a p-value and its implications without truly introducing these ideas formally. The ultimate goal is to have students use computers or even physical simulation to understand uncertainty (such as using a special deck of cards, or even flipping coins) and develop an intuition towards statistical thinking in general.

Ex 18: Discuss the following study and in plain language, report what is stated.

a) Testing the following hypotheses, $H_0 : \mu_D = \mu_C$ vs. $H_a : \mu_D > \mu_C$, subjects in the drug group scored significantly higher ($\bar{x}$ = 23) than did subjects in the control group ($\bar{x}$ = 17). The test statistic of 2.4, gave us a *p*-value of *p* = 0.0135.

The study looked at an experiment with two groups, a control group and a drug group. We will assume the participants were randomly assigned to the groups. The average for the drug group was 23 and the average for the control group was 17. The p-value tells us that the drug group's mean was significantly higher than the control group's mean. When the p-value is small, we reject the null hypothesis (that the two means are the same) and have evidence for the alternative hypothesis. The p-value of .0135 tells us that assuming the two means are equal there is only a 1.35% chance of finding these results or ones where the difference is even more extreme. Since this percentage is small and not very likely to happen, it appears that chance is not at work here and that there truly is a difference between the two means.

b) Note the difference with this statement: Although subjects in the drug group scored higher ($\bar{x}$ = 23) than did subjects in the control group, ($\bar{x}$ = 20), the difference between means was not significant, with a test statistic of 0.98, and a *p*-value of *p* = 0.179.

It would not have been correct to say that there was no difference between the performances of the two groups. There was a difference. It is just that the difference was not large enough to rule out chance as an explanation of the difference. With a p-value of 0.179, which is large, we would say that about 18% of the time we will get a difference this large or larger just by chance assuming the two means are equal. Since this is not that unlikely, we fail to reject the null hypothesis and we do not have evidence of a significant difference between the two groups' means.

Ex 19: Researchers looked at the anxiety score of 10-year old children with alcoholic parents. They compared the data from this sample to the population mean anxiety score for all 10-year old children. The results were reported as follows: The mean score of children of alcoholic parents ($\bar{x}$ = 8.1) was significantly higher than the population mean (μ= 7.0), with a test statistic of z = 2.2, and a *p*-value of *p* = 0.028.

In this study, the researchers took an assumed random sample of 10-year old children with alcoholic parents and through some type of testing, found an anxiety score for each child and then found the average anxiety score for the sample, which was 8.1. Then researchers compared this average to the population mean for all 10-year old children, performed a hypothesis test with the following hypotheses:
$H_0 : \mu = 7$ vs. $H_a : \mu > 7$. **Since the *p*-value is less than 0.05 and is small, the null hypothesis is rejected. It is concluded that the mean anxiety score of 10-year-old children with alcoholic parents is higher than the population mean.**

SAMPLE EXAM QUESTIONS

1. A basketball player makes 2/3 of his free throws. To simulate a single free throw, which of the following assignments of digits to making a free throw are appropriate?

A. 0 and 1 correspond to making the free throw and 2 corresponds to missing the free throw

B. 01, 02, 03, 04, 05, 06, 07, and 08 correspond to making the free throw and 09, 10, 11, and 12 correspond to missing the free throw

C. Both (A) and (B) are correct

D. Neither (A) nor (B) is correct

Ans: C

2. To simulate a single roll of a die, we can use the correspondence 1, 2, 3, 4, 5, and 6 in the table of random numbers. For two consecutive rolls we can use the correspondence

A. 11, 22, 33, 44, 55, 66

B. 11, 12, 13, 14, 15, 16, 21,…26,…61, 62,…66, for 36 possible outcomes

C. Both (A) and (B)

D. Neither (A) nor (B)

Ans: B

3. The General Social Survey (GSS) finds that 28% of the 1500 people interviewed do not approve of capital punishment. The number 28% is

A. A confidence level

B. A parameter

C. A random digit

D. A statistic

Ans: D

4. A candy factory produces chocolate covered in candy shells that come in five different colors. Two types are produced, regular (without peanut) and peanut. Wendy sampled 10 bags of regular chocolate and 10 bags of peanut chocolate. She found that the mean amount of blue pieces in a regular bag was 12.8 and the mean amount of blue pieces in a peanut bag was 9.7. Which statement is true?

A. 10 bags is a parameter

B. 9.7 candies and 12.8 candies are both parameters

C. 9.7 candies and 12.8 candies are both statistics

D. None of these statements are true

Ans: C

5. Which of the following is a correct representation of a proportion?

A. The average amount of hours that high school males spend playing video games

B. A study that found its poll had a margin of error of 3%

C. A study that 75% of all employees spend an hour eating lunch

D. A study that claims that people who regularly use a computer at work have an increased change of developing Carpel Tunnel syndrome

Ans: C

6. A tetrahedron has four triangular faces. Joanne has a die in the shape of a tetrahedron. She rolls the tetrahedron 90 times and records how many times it lands on each side. Which set of results would provide the <u>strongest</u> evidence that Joanne's die is fair?

A. The die lands on the first side 21 times, the second side 19 times, the third side 25 times, and the fourth side 25 times.

B. The die lands on the first side 30 times, the second side 32 times, and the third side 28 times.

C. The die lands on the first side 20 times, the second side 15 times, the third side 33 times, and the fourth side 22 times.

D. The die lands on the first side 23 times, the second side 22 times, the third side 35 times, and the fourth side 10 times.

Ans: A

7. Donna believes that each face of a cube, numbered 1 through 6, is as likely to land face up as any other. To test her hypothesis, she rolls the cube 270 times and records how many times each number lands face up. How many occurrences of each number would provide the strongest evidence of Donna's hypothesis?

A. 270

B. 45

C. 90

D. 42

Ans: B

8. Voters in Jackson County are going to vote on a half-percent sales tax increase to support music in local schools. According to a random survey, 40% plan to vote for the tax and 60% plan to vote against it. The survey's margin of error is ±6%.

Determine whether the survey clearly projects whether the sales tax will pass. Explain your response.

A. The survey does not clearly project whether the sales tax will pass; up to 40% +12% = 52% might vote for the tax and only 60% – 12% = 48% might vote against the tax. The intervals overlap, so the survey does not clearly project the outcome.

B. The survey clearly projects that the sales tax will not pass; 40% ± 6% = 34% to 46% plan to vote for the tax and 60% ± 6% =54% to 66% plan to vote against the tax. The intervals do not overlap, so the survey clearly projects the outcome.

C. The survey clearly projects that the sales tax will pass; 60% ± 6% = 54% to 66% plan to vote for the tax and 40% ± 6% = 34% to 46% plan to vote against the tax. The intervals do not overlap, so the survey clearly projects the outcome.

D. The survey clearly projects that the sales tax will not pass; 40% ± 3% = 37% to 43% plan to vote for the tax and 60% ± 3% = 57% to 63% plan to vote against the tax. The intervals do not overlap, so the survey clearly projects the outcome.

Ans: B

9. An oil company plans to add a chemical to its gasoline to make it burn more cleanly. The company conducts an experiment to see whether adding the chemical affects the gasoline mileage of cars using their gasoline. State the null hypothesis for the experiment.

A. Adding the chemical does not affect gasoline mileage

B. Adding the chemical affect gasoline mileage

C. Adding the chemical decreases gasoline mileage

D. Adding the chemical increases gasoline mileage

Ans: A

10. A study compares how much faster someone runs after drinking a cup of coffee compared to someone who drinks only a cup of water. If the data seems fairly different, what would be the best course of action?

A. Assume there is no difference for people who are unaffected by caffeine

B. No further testing is necessary since there clearly exists a significant difference

C. Drink coffee right before you go running

D. Run a statistical test because it is best not to assume anything

Ans: D

11. The correct way to phrase a null hypothesis for a study where the averages of two sample sets are assumed to be the same is?

A. The significant differences between Sample 1 and Sample 2 are due to chance alone

B. There is no significant difference between the mean of Sample 1 and the mean of Sample 2

C. Sample 1 is assumed to be greater than Sample 2

D. There is a significant difference between the two samples

Ans: B