First Intelligent Online Tutorial System in WPI:

Methods of Tutoring in College Statistics with the ASSISTment System

An Interactive Qualifying Project Report

Submitted to the Faculty
Of the
WORCESTER POLYTECHNIC INSTITUTE
In partial fulfillment of the requirements for the
Degree of Bachelor of Science
By

____

Yurong Mao

____

Alexander Volfson

____

Shengshi Zhao

Date: April 25st, 2008

Approved:

____
Professor Ryum Kim, IQP Advisor

____
Professor Neil Hefferman, IQP Advisor

Abstract

This Interactive Qualifying Project employed computer-based intelligent tutorial system in two Statistics courses, MA2611 and MA2610, for the first time in Worcester Polytechnic Institute. We conducted three randomized experiments to compare the effectiveness of different teaching methods using the online tutoring ASSISTment system. We developed content to compare the effectiveness of ASSISTment over paper-based tutoring and that of hints over worked examples. The study showed that ASSISTment improved student learning and hints were more effective than worked examples.

Acknowledgements

Professor Ryum Kim

Professor Rob Weitz

Professor Neil Heffernan

Criselda Toto

Shane Almeida

Manasi Vartak

Jozsef Patvarczki

Cristina Hefferman

Worcester Polytechnic Institute

Table of Contents

Table of Figures

Table of Tables

1. Introduction

Various projects at Worcester Polytechnic Institute have worked on the development and improvement of the ASSISTment system. In the past, the ASSISTment system had only been used with nearby middle and high school students. This Interactive Qualifying Project is the first use of ASSISTment with college student for the explicit goal of improving teaching methodology. Our goal is first; to develop ASSISTment content for college level statistics teaching; second, to confirm the efficacy of the usage of scaffolding questions in tutoring over paper based worked examples; third, to compare the helpfulness between worked examples and hint messages.

Lacking existing statistics problems in the ASSISTment system, we first constructed individualized, college level, tutoring problems, along with hints and scaffolding problems, for the Worcester Polytechnic Institute MA2611 Applied Statistics 1 course. With the assistance of our reviewing advisors, our group proposed, drafted, and finalized over one hundred problems. To fulfill our second goal, we randomly chose two lab sessions among four and assigned our ASSISTment problems to the chosen sessions. The rest students practiced with same paper-based examples. At the end of the course, we analyzed their quiz scores as well as any feedback and comments submitted by the students. By assigning curriculums containing help in paper-based format to some students and in scaffolding to others and comparing their performance in relevant quizzes we were able to detect whether learning took place in general and whether the learning that took place in one method over the other was statistically significant. After analyzing students' quiz scores, we concluded that ASSISTment did increase learning.

For our third goal, we constructed problems for the course MA2610 and randomly assigned each student ASSISTment problems with either worked examples or a series of hint messages. This time we used the variablization for each problem, resulting in students being assigned one of ten similar but different problems.

ASSISTment is designed to assist teachers to analyze the student performances with better accuracy as well. Teachers have special access to see how each individual student in their class is performing and whether their scores are improving or not. They can easily see the statistics of their student performances by problem or homework. With this data teachers can decide whether they want to review the material again or go on to the next section. It helps the teacher to summarize all of this data by student or class efficiently. Another advantage of the ASSISTment system is that it helps teachers see the improvement of their students easily. It becomes more convenient for teachers to see the progress of students indicated by both numbers and graphs. Teachers can quickly respond according to the students' performance. Is a new teaching method effective? Is the class lecture going too fast? With the data illustrated by the ASSISTment system, these questions can be answered immediately.

This Interactive Qualifying Project is intended to determine how much the student's understanding of statistics may be improved by using the ASSISTment system. To achieve this main purpose, several experiments were designed to conduct at Worcester Polytechnic Institute. The experiment consists of a pre-test and a post-test. We randomized all of the students into two groups. One group received the ASSISTment-based pre-test with hints and scaffolding problems, the other group received a paper-based pre-test with the same content. After the homework, both groups took the same quiz. We analyzed their quiz performance to see whether ASSISTment improved learning better than traditional teaching method. The experiments and details will be explained further in a later section.

2. Background

2.1 The Massachusetts Comprehensive Assessment System (MCAS)

The Massachusetts Comprehensive Assessment System (MCAS) is designed to meet the requirements of the Education Reform Law of 1993. This law specifies that the testing program must: 1. test all public school students in Massachusetts, including students with disabilities and limited English proficient students; 2. measure performance based on the Massachusetts Curriculum Framework learning standards; 3. report on the performance of individual students, schools, and districts. (About MCAS, 2007).

Worcester Polytechnic Institute ASSISTment Interactive Project Groups have built and evaluated a lot of ASSISTment contents, including problem sets and single scaffoldings, based on the MCAC standardized tests for grades 3rd through 10th. There have been many matural reports based on the analysis of this field. Our project, however, is designed for Worcester Polytechnic Institute students as a teaching tool for MA2611/MA2610 (Statistics 1).

2.2 ASSISTment

Limited classroom time available in university, especially Worcester Polytechnic Institute, requires teachers to choose between time spent assisting students' development and time spent assessing their abilities. To help resolve this dilemma, assistance and assessment are integrated in a web-based system called the ASSISTment [1] System that offers instruction to students while providing a more detailed evaluation of their abilities to the teacher than is available under most current approaches. (Neil T. Heffernan, 2006) Traditionally in a statistics class, the instructor focuses on the theory and examples, while students have to work on their own to absorb the material. Within an-hour time limit, it is impossible for the instructor to know whether a student is following or not. Many professors use paper-based homework to evaluate class development and understanding. However, paper-based homework increase the work amount of instructors and also decrease the available time that they could use to prepare for the next class. Also the feedback from students indicated that paper-based homework could not actively interact with a student on a specific question. He or she still had to go to the professor for help. So the homework is really just a way to practice rather than to teach. Now the question becomes: Do we have more effective teaching methods? Yes, we do. The ASSISTment technology provides students with intelligent tutoring assistance while the assessment information is

[1] The term ASSISTment was coined by Kenneth Koedinger and blends Assisting and Assessment.

being collected. ASSISTment is originally constructed by Feng, Heffernan and Koedinger from Worcester Polytechnic Institute Computer Science Department. We could find the introduction of ASSISTment from “Predicting State Test Scores Better with Intelligent Tutoring Systems: Developing Metrics to Measure Assistance Required” by Mingyu Feng, Neil T. Heffernan and Kenneth Koedinger.

> An initial version of the ASSISTment system was created and tested in May, 2004. That version of the system included 40 ASSISTment items. There are now over 1000 ASSISTment items. The key feature of ASSISTment is that they provide instructional assistance in the process of assessing students. The hypothesis is that ASSISTment can do a better job of assessing student knowledge limitations than practice tests or other on-line amount and nature of the assistance that students receive as a way to judge the extent of student knowledge limitations.

There are several advantages of ASSISTment: 1. It is easy to carry out randomized controlled experiments in ASSISTment. 2. The interactive scaffolding questions are well-organized enough to help students with the possible confusions. 3. The pictures in the problem body and the hints can help with the understanding of the theory behind. Being improved all the time, ASSISTment system is now a dynamic system carrying out the randomized variablization feature. Now it allows the instructor to construct random variables for each problem to prevent cheating.

Problems related to the same section are assigned to students in one problem set. For multiple problems in the problem set, the instructor can select the desired problem sequence type. Currently existing section types include “Linear” (problems or sub-sections are presented in linear order), “Random” (problems or sub-sections are presented in a pseudo-random order), and “Experiment” (a single problem or sub-section is selected pseudo-randomly from a list, the others are ignored). (Zachary A. Pardos, 2006)

For each tutoring item, which we call an ASSISTment, is based upon the textbook of the current WPI statistics course. If students get the item correct, they are advanced to the next question. Otherwise, they are provided with a small “tutoring” session, which is the composed of scaffolding questions, where they are asked to answer a few questions that break the problem down into steps. The first scaffolding question appears only if the student gets the item wrong. As long as the student requests for help, including hints and scaffolding questions, the problem will be marked as incorrect on the summary page for the instructor. Students are only marked as correct only if they answer the question correctly on the first attempt.

The summary page allows the instructor to view the development of the students conveniently for the future data analysis. An individual report is generated automatically after a student's completion of the problem set, and the summary report will be automatically updated at the same time.

The summary of the spring 2006 Interactive Qualifying Project experiments described above showed that scaffolding led to higher averages on a post-test, although it was not statistically significant. Here we conducted two experiments using P-test and T-test individually, hoping to get a more significant statistics difference. There purpose of the first experiment is to determine whether ASSISTment is more effective than paper-based materials in terms of teaching methodology. The other experiment is to decide whether hints work better than worked examples in ASSISTment environment. We collect data and student feedback after each experiment to help with the study.

3. Methodology

As described in the introduction, we wanted to establish whether the ASSISTment system could be used to the benefit of college statistics students. To do so we first split students into approximately two equal groups of the students. One group used ASSISTment and the rest received an equivalent packet of information. The student's quiz scores were used to establish the efficiency of the ASSISTment System for college statistics students.

Having established the efficiency of the ASSISTment System, we sought to establish what teaching methodology helps students learn the most. Students who requested assistance on a problem received either a series of hints that guided them through the problem step-by-step or a worked example; one hint that contained a similar problem along with its full solution. Again, paper-based assessments were used to identify learning.

3.1 ASSISTment vs. Paper-Based

First of all we need to introduce our scaffolding system, which played an important role in our first experiment. The idea of scaffolding problems is to break a problem into simple parts to help the student understand the material. A student might have to take several steps to complete a problem, but scaffolding breaks the problem down into manageable parts and walks through the problem with the student. We will take a look at one example here.

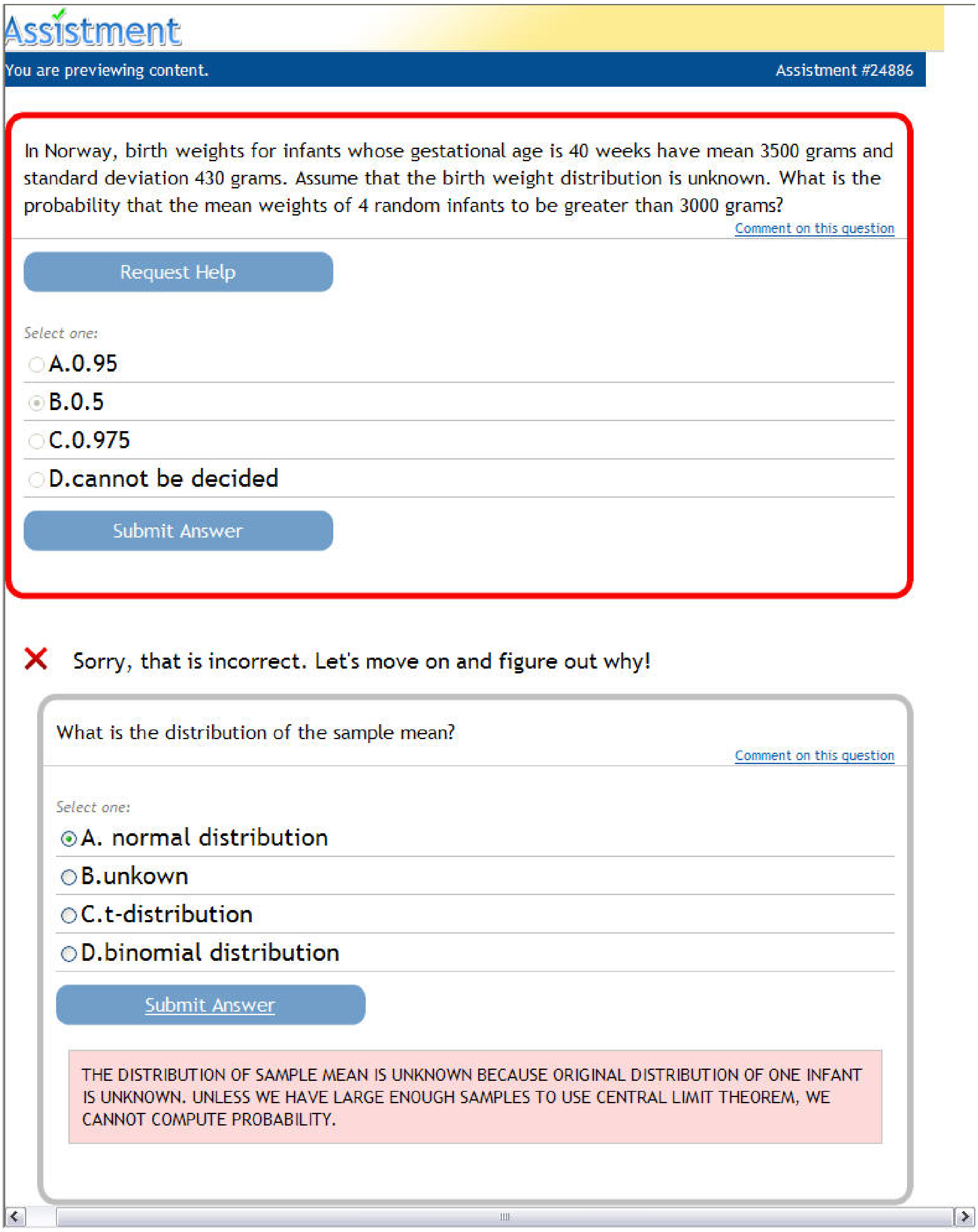


Fig. 3.1 A typical interface of a typical Scaffolding Problem

A student has to do the scaffolding problem if he or she gets a problem wrong. The scaffolding problems cover all the concepts needed in solving the original one. After the student answers all the scaffolding questions correctly, he or she will go back to the original problem and have a chance to do it again.

The ASSISTment System, having been primarily used for teaching MCAS originally, contained no statistics problems. Thus, we began by creating a variety of statistics problems covering the entire curriculum from study design through p-tests. Creating approximately a dozen problems per topic, questions were reviewed for accuracy, clarity, and engagement. With the assistance of our reviewing advisors, over one hundred problems were registered on the ASSISTment System.

Next, two of the four lab sections of students in the Applied Statistics 1 course were randomly to the ASSISTment group; they would do their homework on ASSISTment with scaffolding (required sub-problems that guide the students to the solution), while the other students would have an equivalent paper-based homework, containing the same information. The quiz scores after each homework in addition to a pre- and post-test where recorded for each student.

3.2 Hints vs. Worked Examples

We wanted to address the question of whether students would learn better with step-by-step hints or with a worked example. To do so, we created two versions of some of the homework problems assigned to students. One version contained hints that lead the student through the problem, giving away pieces of the solution sequentially. Often this consisted of three or four hints ending with the answer. The other version contained one big hint that was comprised of a problem that was similar to the main problem as well as an explanation of how a student might solve it.

We generated 10 slightly varying versions of each problem by changing one of the values in the problem within a range. The intent was two-fold; first, we hoped that by making the problem a bit different students would cheat less and secondly, that if they did 'cheat' they would have to explain their methods to do so; resulting in learning. The 62 students of the Applied Statistics for the Life Sciences course were assigned one of ten versions of each problem; the harder of which were supplemented by either hints or worked examples.

We compared hints to worked examples to see which one increase the students' learning. Hints show up every time a student clicks the 'Request Help' button explaining the problem to the student step by step. Worked examples, on the other hand, provide the student with another similar problem and its solution to help his or her understanding. The following pictures demonstrate what a typical problem with hints and the same problem with a worked example look like.

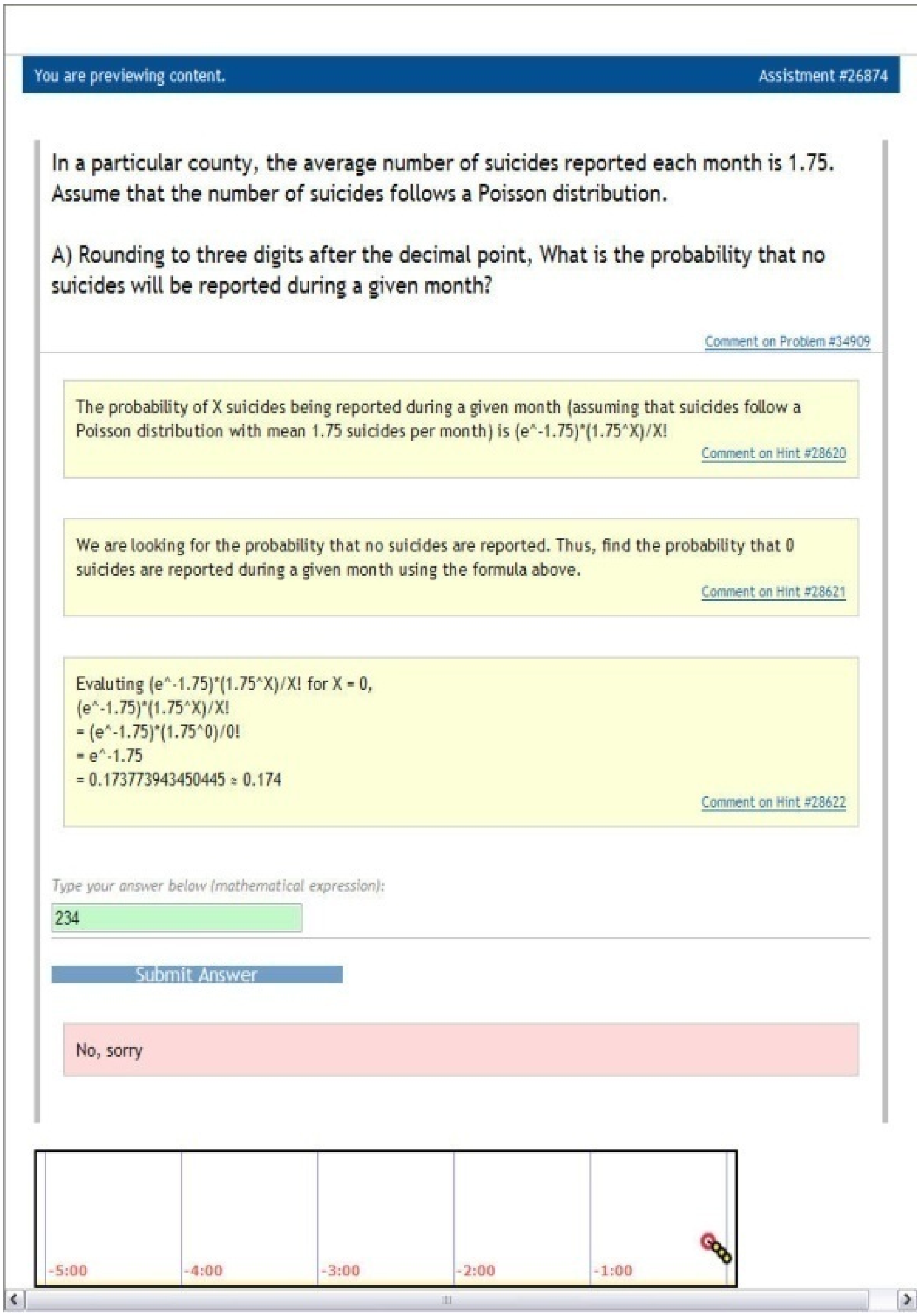


Fig 3.2 Hints Interface

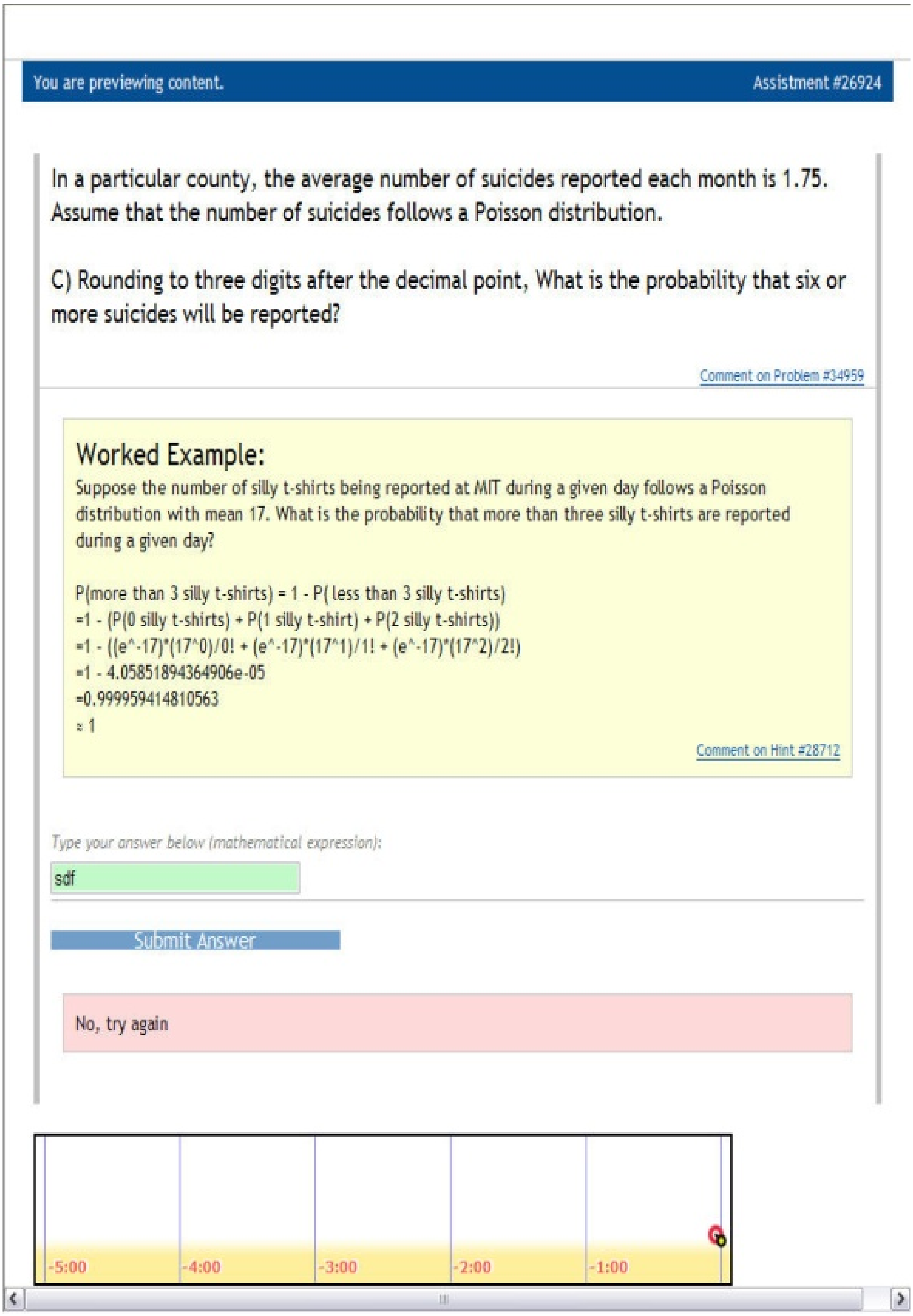


Fig 3.3 Worked Example Interface

Two homework assignments were thusly administered. Student performances on each assignment, as well as the quiz following the first assignment and the mid-term test following the second were recorded for analysis.

4. Trials and Analysis

4.1. A term Analysis on Chapter 3

4.1.1 Summary

On Sep. 12, 2007, we conducted our first scaffolding trial based on the content of *Designing Studies and Obtaining Data* from *Applied Statistics for Engineers and Scientists*. We randomly divided students into two groups, including 56 students in the scaffolding group and 28 students in the pen-paper group. We constructed 4 problems for the ASSISTment tutorial and analyzed the quiz score, comparing to the performances of students using paper-based materials (See Appendix D). In the tutorial, each problem contained 5 scaffolding problems (See Appendix A). After the first trial, we summarized and analyzed the student performances, which would be explained in detail in the later sections. We eliminated one outlier in the scaffolding group because this particular student achieved 100 in the relatively harder pre-test, but did not attend the quiz. This situation was not representative, so we decided to eliminate this outlier. We made our conclusion based on the differences of scores between ASSISTment tutorial and quiz.

4.1.2 Graphs

From the histogram, we could see that the scaffolding group did better than the pen-paper group. There were many more students who attained a score around 80 in the scaffolding group, while the pen-paper group scores were more densely distributed around 70.

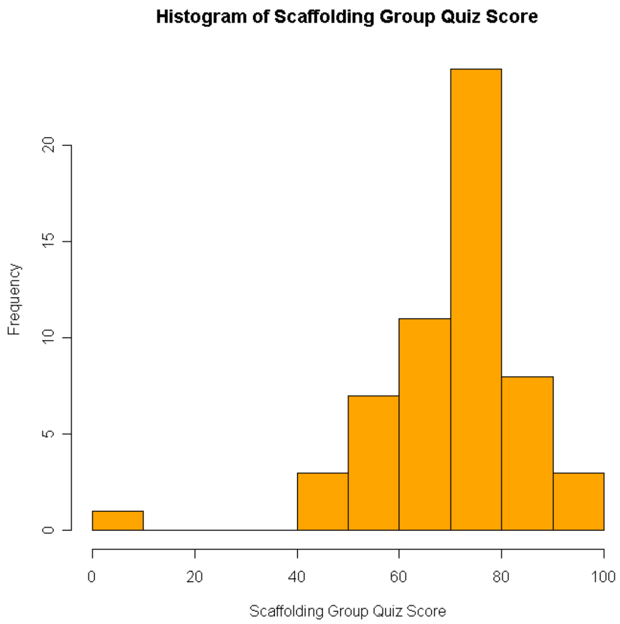


Fig 4.1 Histogram of Scaffolding Group Score

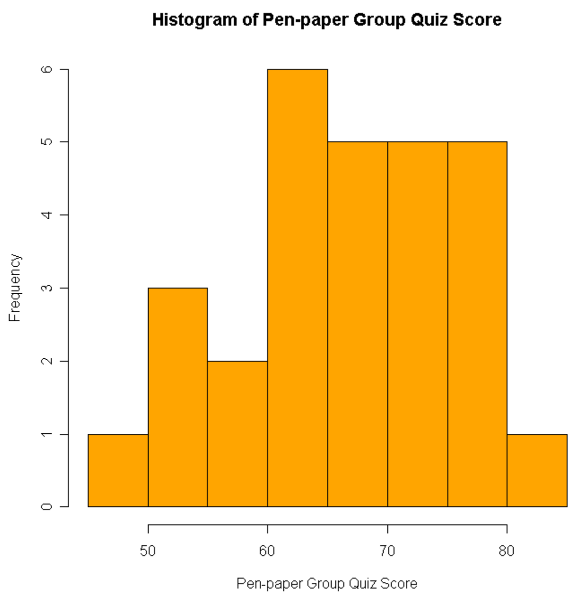


Fig 4.2 Histogram of Pen-paper Group Quiz Score

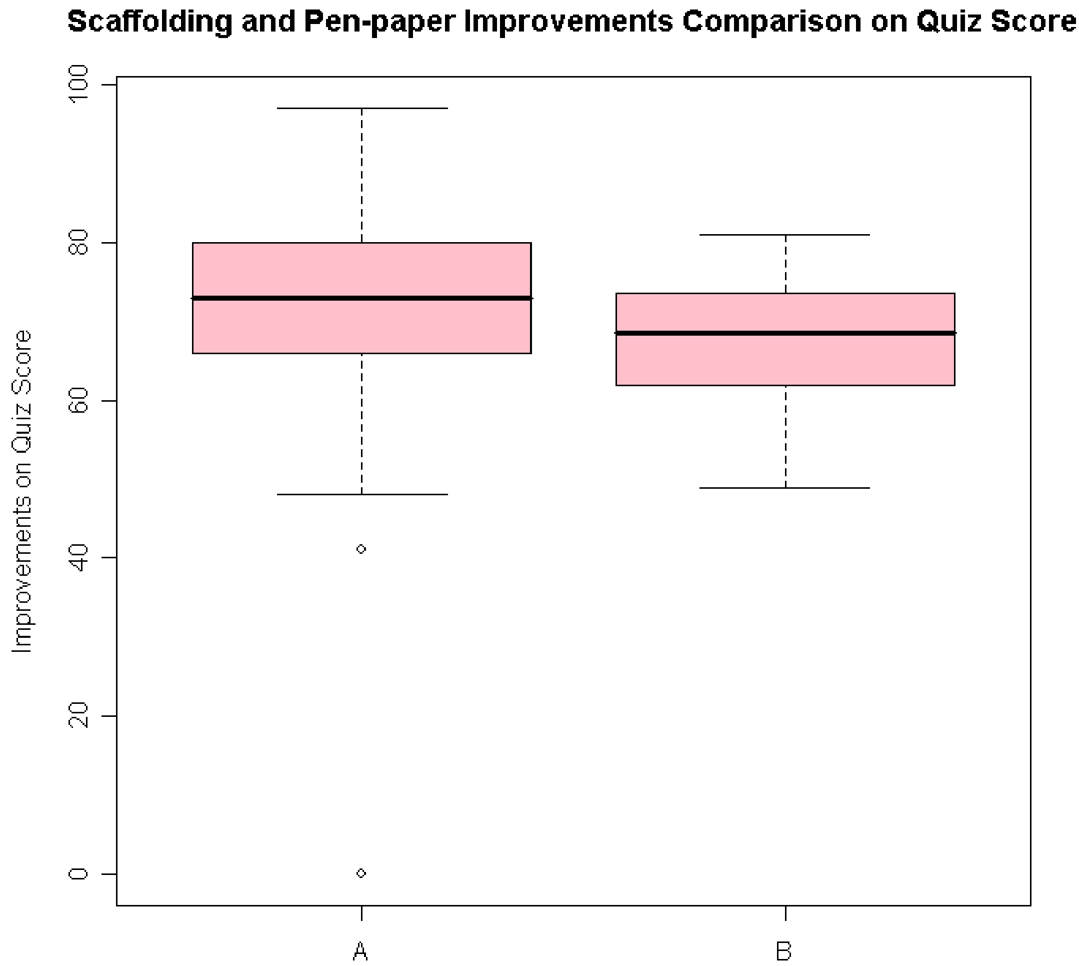


Fig 4.3 Box Plot of Comparison on the quiz Score

In the box plot, the scaffolding group (on the left) had a better mean and median compared with the pen-paper group (on the right). The scaffolding group was distributed mostly above 70, while the pen-paper group was mostly distributed below 80. So from the sharp comparison of the quiz score between two groups, we could conclude that the scaffolding problems helped with the understanding of the materials.

4.1.3 Analysis

Our mean and median for the scaffolding group and the pen-paper group are as follows:

Scaffolding Group					Pen-paper Group				
	Mean	Median	Class Size	Standard Deviation		Mean	Median	Class Size	Standard Deviation
Quiz 3 Score	71.786	73	56	11.54865	Quiz 3 Score	67.214	68.5	28	8.521681

Table 4.1 Analysis with quiz score as post-test

T = (y1 – y2) / ($\sqrt{\frac{S1^2}{N1} + \frac{S2^2}{N2}}$) = 2.0498,

Degree of freedom = $\frac{\frac{S1^2}{N1} + \frac{S2^2}{N2}}{\frac{\left(\frac{S1^2}{N1}\right)^2}{N1-1} + \frac{\left(\frac{S2^2}{N2}\right)^2}{N2-1}}$ =14.12367=15.

To construct a 95% interval, we found the p-value to be 0.029145 < 0.05

So we made a conclusion that the difference between two means was statistically significant.

From the data table (See Appendix C), we could see that the scaffolding group had a better mean and median than the pen-paper study group students. The mean was 3.32 points higher and the median was 4.5 points higher. It proved that the scaffolding problems did help the students understand the material. Students attained better scores after they had gone through scaffolding problems. The quiz was a more reliable measure since each student finished the quiz individually and seriously during the lecture time rather than during the lab time. So it was reasonable to believe that the data that quiz reflected was more trustworthy.

4.1.4 Conclusion

At first, we had an initial ASSISTment-based assessment right after the ASSISTment tutorial online. However, the ASSISTment-based assessment score could not be used during the analysis because students did not treat the post-test as serious as the quiz. So we decided to give up the original test and use the quiz scores as our measurement data. From the quiz scores of the students (See Appendix C), we could conclude that ASSISTment had significantly effect on students' understanding of the course materials. Students who were in the scaffolding group had statistically significantly improvements comparing to the pen-paper group. We could reach the conclusion that ASSISTment made a positively significant impact on the teaching effect.

4.2. A term Analysis on Chapter 4

4.2.1 Summary

Two weeks after the first trial, we conducted our second ASSISTment trial based on Chapter 4. The topic is about statistical model, specifically, central limit theorem. There were 30 scaffolding problems in all on Central Limit Theorem (See Appendix A) to help students with better understanding on the theorem. 17 students in this statistics class were assigned the paper based material (See Appendix E), and 31 students were assigned with 30 scaffolding problems. They had a quiz on the same topic—Central limit theorem. Then we compared the effectiveness of two methods by analyzing their quiz scores.

4.2.2 Graphs

Using the difference score for two groups and R software, we graphed box plot the quiz performance for pen-pencil and scaffolding group. Then we calculate the mean, standard deviation and other estimates to compare the two teaching methods.

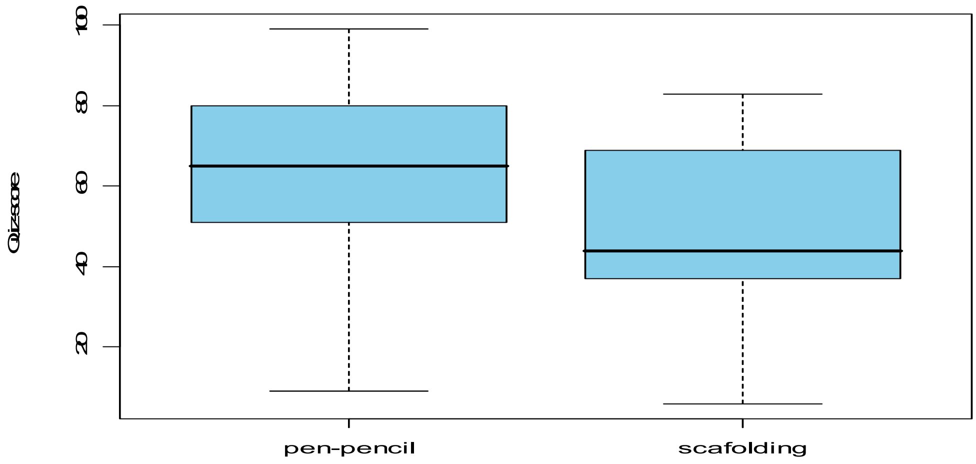


Figure 4.3 Box-plot for quiz score of two groups

The box plot below shows the quiz performance for two groups. The median score of scaffolding group is less than the pen-pencil.

4.2.3 Analysis

The following table summarized two groups' performances for the quiz. On central limit theorem.

Group	Number of Students	Mean	Median	Standard deviation	IQR	0.25th quartile
Pen-pencil group	43	**63.72**	65	22.55	29	51
ASSISTment group	20	**55.3**	50.5	18.57	31	41.25

Table 4.2 Estimations for quiz 4 improvement of both groups

Looking at the mean of two groups' quiz score, pen-pencil group seems to be better than ASSISTment group. Therefore, t test is used to determine whether there is significant difference between two groups' mean score. Using R command of Welch two-sample t-test, we got the result as below.

Mean ASSISTment	Mean Pen-pencil	p-value	95 percent confidence interval:			
55.3	63.72	0.12	lower limit	-2.44	upper limit	19.28

Table 4.4 Confidence Interval for difference of mean quiz score

As p=0.1254 > α=0.05, we cannot reject null hypothesis that the two groups have same mean. The confident interval contains 0, which further proved that there is no significant difference between two groups' mean score. However, the mean score of pen-pencil group is a lot higher than Assistment group's mean score.

4.2.4 Conclusion

Because we cannot reject the null hypothesis, we cannot conclude that pen-pencil group is better than Assistment group. However, the mean performance of two groups showed that pen pencil did better than Assistment. We believe this is because of the relatively poor quality of scaffolding problems for this chapter. The scaffolding problems were highly repetitive; students did not like the content.

4.3 D-term Analysis comparing hints and worked example

For the third trial, 52 students were assigned homework through ASSISTment from March 25th until April 7th of 2008. Students were divided into two groups; those that received hints and those that received worked examples. When students in the worked example group "Requested Help," they received a full solution to a problem that was similar to the one they were working on, but not quite the same. Students in the other group received a series of up to three hints leading them through the solution to their problem. Students who did not attempt all four problems or did not see any hints or worked examples were excluded from the study because they either did not complete all of the questions or they did not receive either treatment. Our analysis is based on the student's performance on one part of one question on the midterm against their performance on four related homework problems (Problem IDs 27032, 27033, 27044, 27045, 27107, 27108, 27109 and 27110 in the Appendix A).

Both groups of students were assigned homework on the topic of normal probability computation in ASSISTment. These problems are created by us using ASSISTment variablization. The same problem would end up with different numbers in the problem body, which effectively prevented students from cheating on the homework. Variablization also helped consolidate the results when analyzing whether the hints or worked example method improved students' learning the most.

There were four ASSISTment problems assigned as homework to students about normal probability computation, but the corresponding midterm question requires both central limit theorem and normal probability computation. Students were randomly assigned the questions. Some of them would get problem with hints showing up after they clicked on "Request Help". Others were assigned worked example problems, which were similar problems to the main question, but provided with solution process. Though 52 students were assigned homework, some had to be excluded from the study because they did not attempt all of the problems. The chart below shows the breakdown of the students by the number of problems they tried:

Number of Problems Tried	***Hints***	***Worked Example***
0	6	
1	1	2
2	2	2
3	0	0
4	22	17
Total	52	

Table 4.4 Number of problems the students attempted

The highest score that a student could get on the midterm problem on normal probability computation was 8 points. If he or she received 0 to 4 points, he or she would be marked as 0 to represent failure, while 5 to 8 points would be denoted as a success and was assigned value of 1. The two-way contingency table of the students follows:

	Hints	***Worked Example***
Correct	11	3
Wrong	6 (64.7%)	10 (23.1)%
Total	17	13

Table 4.5 Student performance by treatment

As ASSISTment automatically records answers when students were doing each problem, the data was easy to access. We analyzed the score on one of the midterm problems for the two groups. The students who got all homework questions right without going over the hints or worked example are not included because they did not receive any tutoring. Students who either went over at least one hint or one worked example problem were included. Also, student must have tried all the homework questions related to the normal probability computation to be included. The table below summarizes the data:

Hints Group			***Worked Example Group***		
ID	***Midterm***	***Success***	***ID***	***Midterm***	***Success***
1	8	1	**1**	8	1
2	0	0	**2**	8	1
3	8	1	**3**	0	0
4	8	1	**4**	3	0
5	1	0	**5**	0	0
6	8	1	**6**	3	0
7	4	0	**7**	2	0

8	8	1	**8**	2	0
9	8	1	**9**	0	0
10	5	1	**10**	0	0
11	8	1	**11**	0	0
12	0	0	**12**	0	0
13	8	1	**13**	8	1
14	0	0			
15	6	1			
16	8	1			
17	4	0			
Total		11			3

Table 4.6 Scores of the two groups

Using these data, the result is following:

Prop. Hint	Prop. Worked	p-value	95 percent confidence interval:			
64.70%	23.10%	0.05802	lower limit	0.026	upper limit	0.807

Table 4.7 Statistical Summary

The two sample proportion test for hints and worked group gave the 95% percent confidence interval of (0.026, 0.0807). As 0 was not contained in the interval, there was significant difference between the proportions of students in hint group who got higher than 4 points and worked example group. Furthermore, the proportion estimate for the hint group students was higher than worked example students. The reason that students preferred hints is because worked examples were quite long and indirect to the original question. Students had the feelings that these examples were almost irrelevant to the main question and did not bother to look at them at all. That was the reason why worked examples did not improve the students' understanding of the material as well as hints did, which were short and penetrated.

Therefore, hints group performed better than worked example group on the normal probability computation problems. In other words, hints work better as a direct and effective tutoring method compared to worked examples. Students are more likely to grasp the concept and solve statistical computation problem with hints provided.

5. Results

For the first time in WPI, we initiated the electronic tutoring of statistics that will be improved in future years. In our first trial, we determined that ASSISTment did improve learning, compared to the typical paper-and-pencil method, in the chapter on Study Design (increase of 6.8%; p-value of 0.03), though the results of our second trial for Statistical Modeling were less conclusive (decrease of 4.8%; p-value of 0.12). In our third trial, we found that step-by-step hints are better than providing a worked example (increase of 41.6%,

p-value of 0.058). We created much content (see Appendix A): 33 problems with scaffolding and 14 in both hints and worked example versions. We also initiated variablization; we worked through some bugs and created 8 variablized problems.

6. Discussion

6.1 Systemic Error

Statistics is a course that is recommended or suggested background for many courses of study. As such, the students' mathematical background varies greatly; not only from student to student, but class section to class section, and year to year. This heterogeneity, had the potential to skew our results if, for example, all of the students with a passion for numbers happened to be grouped together. We avoided this as much as possible by assigning problem sets to student uniformly at random during our second trial and assigning problem set types to class sections at random during the first. As in the second trial, it would have been ideal to assign ASSISTment and paper-based completely at random during the first trial, however since student had shared time to work on these problems with Teaching Assistant help, this would have caused confusion.

6.2 ASSISTment

Another factor that may have affected our results was the use of the ASSISTment system itself. Since it was being developed while we worked on problems, our team faced numerous time consuming challenges. The content-creation side of the ASSISTment system is not nearly as easy to use as the student side. The content creation lacks common features from “Save As…” to “Print,” an inability to search through the plethora of old problems and the general inflexibility of editing required the group to spend much of its time and effort wrestling with the system. For the first trial, for example, future ASSISTment problems were first written in text documents and then copy-pasted into corresponding ASSISTment fields.

As we added problems, new features were being added to the ASSISTment system, with mixed results. For the second trial, both the ability to “variablize” problems and access to the R software environment for statistics were added. With the variablization, variations of problems could be generated easily, reducing the ease with which students could cheat and in case students did collaborate, forcing them to explain their steps; while with the R software, calculations could be done inside of problems automatically. Since security holes had to be created to make these features work, we had to interface with a developer on the ASSISTment side to help us make our variablized problems accessible to students. When this became a bottleneck, the quality of the homework material suffered because there was no room for error and no time for feedback. Coincidentally, the developer’s work also involved rote copy-paste.

Despite these flaws with ASSISTment, our process could have been better as well. For the first homework assignment in the second trial, a miscommunication on the team resulted in less problems being created than expected. For the second homework assignment, we made sure everyone knew their assigned tasks and avoided this issue. Because of time constraints only some problems were done with in both hints and worked example variations. Unfortunately, despite the production of nearly two-dozen problems, an oversight resulted in only four homework problems being applicable to our study, and, even worse, only one part of one problem on the midterm exam corresponded to these homework problems. This resulted in a much smaller data set than we expected.

6.3 Conclusions

From the first trial, students who used ASSISTment learned more because of the personalization of the system. Hints, messages and scaffolding problems that the ASSISTment group received corresponded directly a student's task at hand. The paper-based group, however, would have had to read through an electronic packet of information to find the relevant parts instead. This could be distracting and lead to more mistakes if the wrong section was identified as relevant.

From the third trial, students who received hints learned more for similar reasons. Both the hints and the worked example were related to the problem with which the student was currently struggling. Thus, the student learned strategies with which to approach the problem. Those students who received hints instead of a worked example, however, saw the strategy broken down step-by-step and may have even completed the problem seeing only the first hint to help them get started and finishing the rest of the problem on their own. With a worked example, this is not entirely the same; the wording and the context may be different, but students may simply speed read the equations instead of analyzing the fine points of the solution. Some students commented that they did not feel that the worked examples were relevant to their homework problem. However, this claim may have come from a simple lack of motivation on the part of the student to really understand the worked example.

Despite the challenges we faced, we did promote student learning. Where ASSISTment had no statistics problems, we initiated the creation of course work for years to come. We also pioneered the use of variablization in the ASSISTment system. When we compared ASSISTment to paper-based assignments, students were better prepared for conceptual material corresponding to study design and statistical modeling. In the second trial, we compared the step-by-step hints methodology to the worked example methodology. We observed significant learning in the hints group despite a smaller than expected data set.

Assistment

Assistment #25980

You are previewing content.
Please select the correct order of the letters corresponding to the study type (as indicated in the graphic at the bottom) which applies to each of the following three problems, in order. Do not calculate anything. Simply identify the study type of each problem with a letter A-G.

1. One thousand students were given a standardized English test and a standardized math test. Twenty-two students were randomly selected from a population of 1000 students, such that each student had an equal probability of being selected (simple random sampling). Test results are summarized below.

Find the 90% confidence interval for the mean difference between student scores of the math and English tests. Assume that the mean differences are approximately normally distributed.

Student	English	Math	Difference, d	$(d_i - d)^2$
1	95	90	5	16
2	89	85	4	9
3	76	73	3	4
4	92	90	2	1
5	91	90	1	0
6	53	53	0	1
7	67	68	-1	4
8	88	90	-2	9
9	75	78	-3	16
10	85	89	-4	25
11	90	95	-5	36
12	85	83	2	1
13	87	83	4	9
14	85	83	2	1
15	85	82	3	4
16	68	65	3	4
17	81	79	2	1
18	84	83	1	0
19	71	60	11	100
20	46	47	-1	4
21	75	77	-2	9
22	80	83	-3	16

2. An airline wants to evaluate the depth perception of its older pilots. A random sample of 14 airline pilots over the age of fifty are asked to judge the distance between two markers placed 20 feet apart at the opposite end of the laboratory. The sample data listed here are the pilots' error (recorded in the feet) in judging their distance to the markers.

2.7 2.4 1.9 2.6 2.4 1.9 2.3

2.2 2.5 2.3 1.8 2.5 2.0 2.2

Use the sample data to place a 95% confidence interval on μ, the average error in depth perception for the company's pilots over the age of fifty.

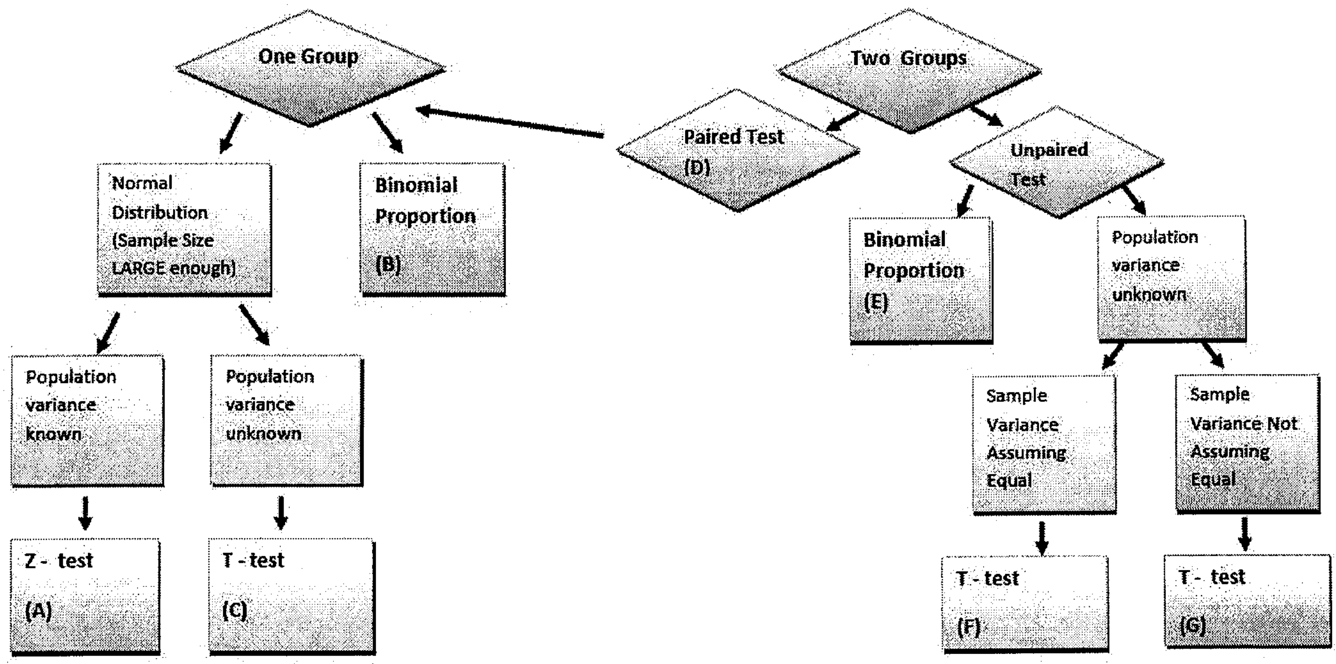
3. An experiment was conducted to evaluate the effectiveness of a treatment for tapeworms in the stomachs of sheep. Twenty-four worm-infected lambs of approximately the same age and health was randomly divided into two groups. Twelve of the lambs were treated with a new drug and the remaining twelve were left untreated. After a 6-month period, the lambs were slaughtered and the following worm counts were recorded:

Drug-Treated Sheep	18	43	28	50	16	32	13	35	38	33	6	7

Appendix A

Untreated Sheep	40	54	26	63	21	37	39	23	48	58	28	39

Place a 95% confidence interval to assess the size of the difference in the two means, assuming that the variances are of a similar size.



Comment on this question
Request Help

Select one:

- ○ CGB
- ○ DCF
- ○ AFB
- ○ CBF
- ○ DAE

Submit Answer

Assistment

Assistment #25981

You are previewing content.
Please select the correct order of the letters corresponding to the study type (as indicated in the graphic at the bottom) which applies to each of the following three problems, in order. Do not calculate anything. Simply identify the study type of each problem with a letter A-G.

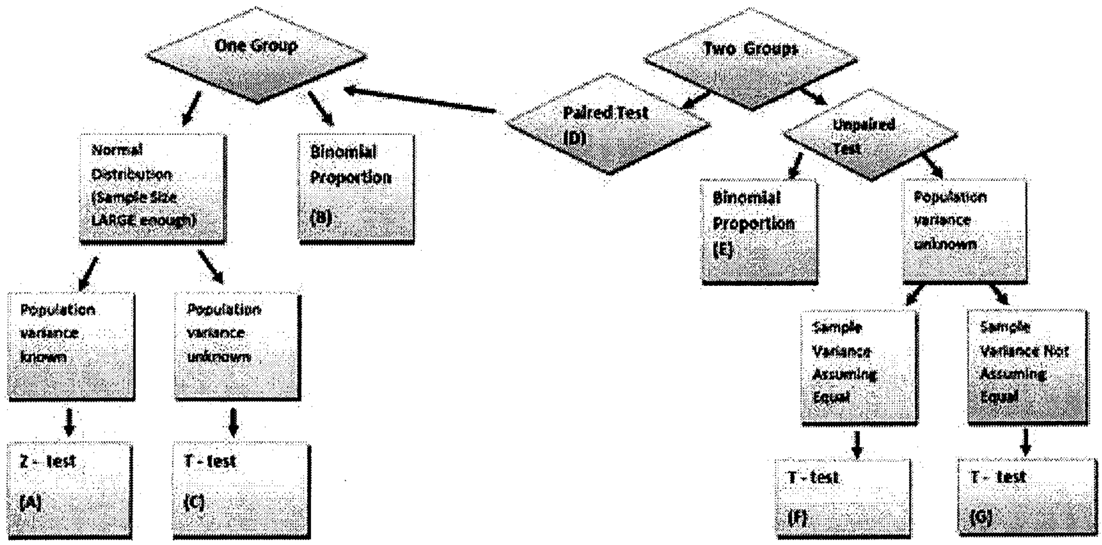
1. Researchers hypothesize that oil spill sites suffer a decrease in plant density. They select a random sample of oil spill sites and control sites on which they measure the plant density. Use the following values (which summarize the the sample data) to construct an appropriate 95% confidence interval for $\mu 1$-$\mu 2$, the difference between the average plant desity of control and oil spill sites.

	Control Plots	Oil Spill Plots
N	40	40
$\hat{Y}$ (Sample Mean)	38.48	26.93
SD	16.37	9.88

2. Suppose a TV station conducts a survey to determine whether the public supports or opposes a new city ordinance. Surveying 900 city residents, 340 support the ordinance and 560 oppose it. What is a 95% confidence interval for the proportion of the city residents that support the proposed ordinance?

3. A forester wishes to estimate the average number of "count trees" per acre (trees larger than a specified size) on a 2000-acre plantation. She can then use this information to determine the total timber volume for trees in the plantation. A random sample of n = 50 1-acre plots is selected and examined. The average (mean) number of "count trees" per acre is found to be 27.3, with a standard deviation of 12.1. Use this information to construct a 99% confidence interval for μ, the mean number of count trees per acre for the entire plantation.

Appendix A



Comment on this question

Request Help

Select one:

- ◯ FEC
- ◯ GDC
- ◯ BGD
- ◯ GBA
- ◯ AED
- ◯ BAG
- ◯ GCA

Submit Answer

Assistment

Assistment #25982

You are previewing content.
Suppose a coach wants to know how fast his six atheletes can run. When he recorded the time that each of them took to finish 400 meters, which were 102.5, 101.7, 103.1, 100.9, 100.5, and 102.2 (seconds). He calculates the mean to be 101.82. What is the confidence interval for the mean time to be at a 95% confidence level?

Comment on this question

Request Help

Select one:

- ○(101.01, 102.63)
- ○(100.86, 102.78)
- ○(99.86,103.78)
- ○(101.57, 102.07)

Submit Answer

Assistment

Assistment #25984

You are previewing content.

1000 randomly selected Americans were asked if they believed that the annual salary for biology engineers would increase in the following year. 600 said yes. Construct a 95% confidence interval for the proportion of Americans who believe that the annual salary for biology engineers will increase.

Comment on this question

Request Help

Select one:

- ○ (42%, 78%)
- ○ (54%, 66%)
- ○ (56%, 64%)
- ○ (57%, 63%)
- ○ (55%, 65%)

Submit Answer

Assistment

Assistment #25989

You are previewing content.

Suppose that you were interested in the average force that a new machine generates. You wanted to find a 95% confidence interval with a margin of error of 0.5 based on all the existing data assuming that the standard deviation is 10. What is the least number of machine runs required?

Comment on this question

Request Help

Select one:

- ○ 1536
- ○ 427
- ○ 399
- ○ 572
- ○ 651

Submit Answer

Assistment

Assistment #25990

You are previewing content.

A Subaru dealer wants a 90% confidence interval of the mean age of his customers for advertising purposes. He wants the margin of error to be 3 years. Assume that the standard deviation of the customers' age is 13. What is the least number of customers he needs to survey?

Comment on this question

Request Help

Select one:

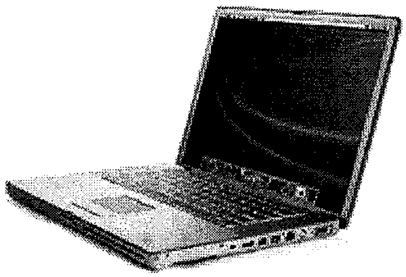
- ○ 51
- ○ 49
- ○ 61
- ○ 57
- ○ 52

Submit Answer

Assistment

Assistment #25991

You are previewing content.



Suppose that you are working in the Apple testing center. You are to test whether smashing a Macbook will damage it, while minimizing the numbers of laptops you need to break. If you want a 90% confidence interval for this test, with a margin of error of 4%, What's the least number of laptops you should smash? (This could be expensive!!!!!)

Comment on this question

Request Help

Select one:

- ○ 125
- ○ 886
- ○ 671
- ○ 426
- ○ 1028

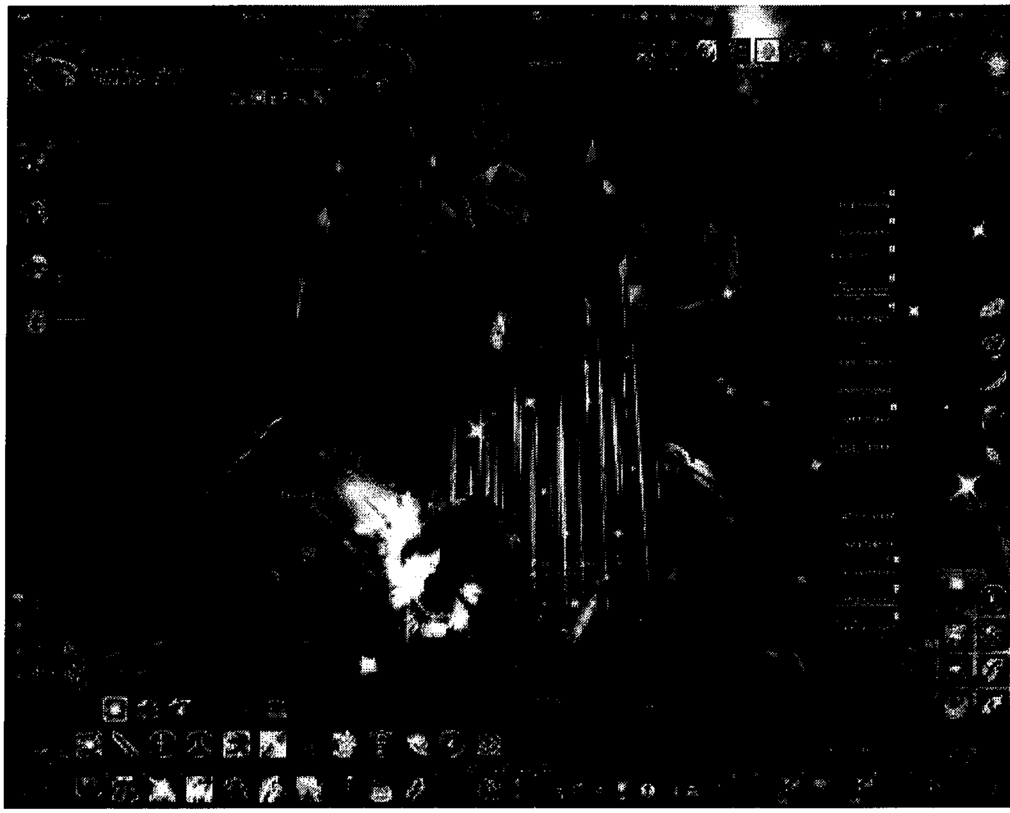
Submit Answer

Appendix A

Assistment

Assistment #25992

You are previewing content.



Two teams are competing in World of Warcraft. One team is formed by all the WPI ECE girls and the other by all of the WPI Math professors. We randomly select 30 girls and 35 professors out of the two teams. The mean number of soldiers generated in the first 15 minutes by the selected girls was 15.6 with a standard deviation of 2.8. The mean number of soldiers generated in the first 15 minutes by the selected professors was 14.3 with a of 9.1.

Construct a 95% confidence interval for the difference between the means.

Comment on this question

Request Help

Select one:

- ○ (2.14, 4.74)
- ○ (1.58, 3.62)
- ○ (3.27, 4.55)
- ○ (-1.33, 4.58)
- ○ (-2.14, 4.74)

Submit Answer

Assistment

Assistment #25993

You are previewing content.

One thousand students were given a standardized English test and a standardized math test. Twenty-two students were randomly selected from a population of 1000 students, such that each student had an equal probability of being selected (simple random sampling). Test results are summarized below.

Find the 90% confidence interval for the mean difference between student scores of the math and English tests. Assume that the mean differences are approximately normally distributed.

Student	English	Math	Difference, d	$(d_i - d)^2$
1	95	90	5	16
2	89	85	4	9
3	76	73	3	4
4	92	90	2	1
5	91	90	1	0
6	53	53	0	1
7	67	68	-1	4
8	88	90	-2	9
9	75	78	-3	16
10	85	89	-4	25
11	90	95	-5	36
12	85	83	2	1
13	87	83	4	9
14	85	83	2	1
15	85	82	3	4
16	68	65	3	4
17	81	79	2	1
18	84	83	1	0
19	71	60	11	100
20	46	47	-1	4
21	75	77	-2	9
22	80	83	-3	16

Comment on this question

Request Help

Select one:

- ○ (0.3, 1.3)
- ○ (1, 2.3)
- ○ (-0.3, 2.3)
- ○ (0.3, 2.3)
- ○ (-0.3, 1.3)

Submit Answer

Assistment

Assistment #25994

You are previewing content.

An experiment was conducted to evaluate the effectiveness of a treatment for tapeworms in the stomachs of sheep. Twenty-four worm-infected lambs of approximately the same age and health was randomly divided into two groups. Twelve of the lambs were treated with a new drug and the remaining twelve were left untreated. After a 6-month period, the lambs were slaughtered and the following worm counts were recorded:

Drug-Treated Sheep	18	43	28	50	16	32	13	35	38	33	6	7
Untreated Sheep	40	54	26	63	21	37	39	23	48	58	28	39

Place a 95% confidence interval to assess the size of the difference in the two means, assuming that the variances are of a similar size.

Save one decimal place

Comment on this question

Request Help

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #25995

You are previewing content.

Insurance adjusters are concerned about the high estimates they are receiving for auto repairs from garage 1 as compared to garage 2. To verify their suspicions, each of 15 cars recently involved in an accident was taken to both garages for separate estimates of repair costs. The estimates from the two garages are given below. What is the 95% confidence interval of the cost difference of the two garages, to two decimal places?

Car	Garage 1	Garage 2
1	17.6	17.3
2	20.2	19.1
3	19.5	18.4
4	11.3	11.5
5	13.0	12.7
6	16.3	15.8
7	15.3	14.9
8	16.2	15.3
9	12.2	12.0
10	14.8	14.2
11	21.3	21.0
12	22.1	21.0
13	16.9	16.1
14	17.6	16.7
15	18.4	17.5
Total	Mean = 16.85	Mean = 16.23
	Sample SD = 3.20	Sample SD = 2.94

Comment on this question

Request Help

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #25996

You are previewing content.
A school district decided that the number of students attending their high school was nearly unmanageable, so they decided to split into two districts, with District 1 students going to the old high school and District 2 students going to a newly constructed building. A group of parents became concerned with how the two districts were constructed relative to income levels. A study was thus conducted to determine whether persons in District 1 have a different mean income from those in District 2. A random sample of 20 homeowners was taken in District 1. Although 20 homeowners were to be interviewed in District 2 also, one person refused to provide the information requested, even though the researcher promised to keep the interview confidential. Thus, only 19 observations were obtained from District 2. The data, recorded in thousands of dollars, produced sample means and variances as shown below. Use these data to construct a 95% confidence interval for μ1-μ2, to two decimal places.

	District 1	District 2
Sample Size	20	19
Sample Mean	18.27	16.78
Sample Variance	8.74	6.58

Comment on this question

Request Help

Type your answer below:

•

Submit Answer

Assistment

Assistment #24668

You are previewing content.

Dr. Thompson and his research group are interested in the effect of vitamin C on preventing flu infection. He was authorized to collect information regarding vitamin C intake history of all patients who were admitted in several hospitals with flu infection. He than collected questionnaires from additional 1000 volunteers who were screened to make sure that none of them had flu symptom this year. The volunteers were asked of their regular vitamin C intake.

What is the design type of the study?

Comment on this question

Select one:

- ◯ A. Completely Randomized Design
- ◯ B. Randomized Completely Block Design
- ◯ C. Retrospective Observational Study
- ◯ D. Prospective Observational Study
- ◯ E. Sampling Survey

Submit Answer

Assistment

Assistment #24671

You are previewing content.
Dr. Weekes's group is also interested in the effect of vitamin C on preventing flu infection. She collected 2000 volunteers. They reported whether they regularly take vitamin C or not. These volunteers were followed for one year to see if they had a flu infection.

What type of study is this one?

Comment on this question

Select one:

- ◯ A. Completely Randomized Design
- ◯ B. Randomized Completely Block Design
- ◯ C. Retrospective Study
- ◯ D. Prospective Study
- ◯ E. Sampling Survey

Submit Answer

Assistment

Assistment #24672

You are previewing content.

Dr. Park's group is also interested in the effect of vitamin C on preventing flu infection. She advertised to give free vitamin C to 10,000 adult volunteers who were screened to make sure they have not already taken vitamin C regularly. For this three-month study, volunteers who received flu vaccination in 2006 were randomly divided into two groups. One group was given vitamin C and the other group received placebo. Volunteers who have not received flu vaccination this year were similarly divided into two groups: one was given vitamin C and the other placebo.

What type of study is this one?

Comment on this question

Select one:

- ◯ A. Completely Randomized Design
- ◯ B. Randomized Completely Block Design
- ◯ C. Retrospective Observational Study
- ◯ D. Prospective Observational Study
- ◯ E. Sampling Survey

Submit Answer

Assistment

Assistment #24810

You are previewing content.

You will be given link to the solutions after you solve these problems. After you study the solutions, you will be asked to solve another set of problems.

These problems will be similar to the quiz problems on this Friday. If you are well aware of the contents in chapter 3, you may need very little time to complete both tests. Please leave the classroom quietly so that you don't disturb your classmates who are still learning.

Now let's solve the problems:

A manufacturer of roofing shingles wants to compare the performance of shingles with two different types of backings in field tests. To do so, they randomly select 30 communities around the county. In each, they randomly select a single-family house among those volunteered by their owners in response to an ad for a "free roof." They randomly select half the houses to receive one type of shingle and roof the rest with the second type. Various measures of the condition of each roof are obtained over a period of years.

What type of study is this one?

Comment on this question

Select one:

- ○ A Completely Randomized Design
- ○ B. Randomized Completely Block Design
- ○ C. Retrospective Observational Study
- ○ D. Prospective Observational Study
- ○ E. Sampling Survey

Submit Answer

Assistment

Assistment #24812

You are previewing content.
Consider the previous problem: another manufacturer is interested in the comparison of two types of shingles. This new manufacturer suspects that the annual precipitation affects the condition of roofs. So he divides the county into three regions: high precipitation, moderate precipitation, and low precipitation area. Within each of three regions, the manufacturer randomly assigns one type of shingle to 5 houses and another type to another 5 houses.

What type of study is this one?

Comment on this question

Select one:

- ○ A Completely Randomized Design
- ◉ B. Randomized Completely Block Design
- ○ C. Retrospective Study
- ○ D. Prospective Study
- ○ E. Sampling Survey

Submit Answer

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24813

You are previewing content.
Consider the previous problem again: another manufacturer is interested in the comparison of two types of shingles. However, this manufacturer does not have time or finance to assign free roof. Instead, she found 100 houses with their roof life span longer than 10 years, and also found 40 houses with their roof life span less than 3 years. Within the long-lasting roofs, 80% had type A shingles and within short-lived roofs, 30% had type B shingles.

What type of study is this one?

Comment on this question

Select one:

- ○ A. Completely Randomized Design
- ○ B. Randomized Completely Block Design
- ◉ C. Retrospective Study
- ○ D. Prospective Observational Study
- ○ E. Sampling Survey

Submit Answer
That is not correct, try again.

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24814

You are previewing content.
Consider the previous problem again: another manufacturer first identified 1000 houses with shingle type A and 500 houses with shingle type B in the county. Among the 1000 houses, 20% had life span shorter than 3 years, and among 500 houses, 50% had life span shorter than 3 years.

What type of study is this one?

Comment on this question

Select one:

- ○ A Completely Randomized Design
- ○ A Completely Randomized Design
- ○ B. Randomized Completely Block Design
- ○ C. Retrospective Observational Study
- ○ D. Prospective Study
- ○ E. Sampling Survey

Submit Answer

Assistment

Assistment #24886

You are previewing content.
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 4 random infants to be greater than 3000 grams?

Comment on this question

Request Help

Select one:

- ◯ A.0.95
- ◯ B.0.5
- ◯ C.0.975
- ◯ D.cannot be decided

Submit Answer

Assistment

Assistment #24886

You are previewing content.
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 4 random infants to be greater than 3000 grams?

Comment on this question

Request Help

Select one:

- A.0.95
- B.0.5
- C.0.975
- D.cannot be decided

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
What is the distribution of the sample mean?

Comment on this question

Select one:

- A. normal distribution
- B.unkown
- C.t-distribution
- D.binomial distribution

Submit Answer
THE DISTRIBUTION OF SAMPLE MEAN IS UNKNOWN BECAUSE ORIGINAL DISTRIBUTION OF ONE INFANT IS UNKNOWN. UNLESS WE HAVE LARGE ENOUGH SAMPLES TO USE CENTRAL LIMIT THEOREM, WE CANNOT COMPUTE PROBABILITY.

Correct!
What is the probability that the sample mean is greater than 3000 grams?

Comment on this question

Select one:

- A.0.95
- B.0.5
- C.0.975
- D.cannot be decided

Submit Answer
WE CAN'T SOLVE THIS BECAUSE WE DO NOT KNOW THE DISTRIBUTION.

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24892

You are previewing content.
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 40 random infants to be greater than 3400 grams?

Comment on this question

Request Help

Select one:

- A.0.889
- B.0.899
- C.0.909
- D.0.919
- E.0.929

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
What is the distribution of the sample mean?

Comment on this question

Select one:

- A. normal distribution
- B. unknown
- C. t-distribution
- D. binomial distribution

Submit Answer

Correct!
Now, let's review how to compute normal probability. What is the probability that a normal random variable with mean 100 and standard devation 25 to be greater than 125? You need to use Z-transformation and use R command pnorm().

Comment on this question

Select one:

- A. 0.16
- B.0.17
- C.0.18
- D.0.19

Appendix A

Submit Answer

Correct!
As the mean of the birth weight of each infant is 3500 grams, what's mean of the sample mean of 40 infants' birth weights?

Comment on this question

Select one:

- ◯ A.3400
- ◉ B.3500
- ◯ C.3600
- ◯ D.3700

Submit Answer

The mean of the sample mean is identical to the mean of each random variable. Here, each random variable (birth weight of each infant) has mean 3500 grams

Correct!
Each of x1, x2, x3,...,xn has standard deviation δ. Then, what's the standard deviation of the sample mean?

Comment on this question

Select one:

- ◉ A δ/sqrt(n)
- ◯ B δ square
- ◯ C δ
- ◯ D δ*n

Submit Answer

Correct!
As the birth weight of any infants has standard deviation of birth weight at 430 grams, what is the standard deviation of the sample mean of 40 infants' birth weights?

Comment on this question

Select one:

- ◯ A.65.45
- ◯ B.66.58
- ◯ C.68.66
- ◉ D.67.99

Submit Answer
we just apply the property in the previous problem and plug in the number into the formula: δ/sqrt(n). Here, δ=430 grams, n=25

Correct!
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 40 random infants to be greater than 3400 grams?
Comment on this question

Select one:

- ○ A.0.889
- ○ B.0.899
- ○ C.0.909
- ○ D.0.919
- ◉ E.0.929

Submit Answer
We know that the sample mean has 1) normal distribution, 2) mean 3500, and 3) standard deviation 67.99. So the problem turns into regular normal probability computation. P(sample mean > 3400) = 1-P(sample mean < 3400) = 1- P(Z < (3400-3500)/67.99) This can be solved in R by the following command: 1-pnorm((3400-3500)/67.99)

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24894

You are previewing content.
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is normal. What is the probability that the mean weights of 4 random infants to be greater than 3000 grams?

Comment on this question

Request Help

Select one:

- A 0.99
- B 0.98
- C 0.97
- D 0.96
- E 0.95

Submit Answer

Correct!
You are done with this problem!

Comment on this problem

Go to next problem

Assistment

Assistment #24898

You are previewing content.
The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the battery life has a normal distribution. Suppose there are 5 batteries on hand. What is the probability that the 5 batteries are used up in less than 4000 hours?

Comment on this question
Request Help

Select one:

- A 0.95
- B 0.975
- C 0.5
- D unknown

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
"What is the probability that the 5 batteries are used up in less than 4000 hours?"

Can you restate this problem in terms of sample mean of battery life, not of the total battery life?

Comment on this question

Select one:

- A. What is the probability that the mean life of 5 batteries are less than 800 hours?
- B. What is the probability that the mean life of 5 batteries are less than 4000 hours?
- C. What is the probability that the mean life of 5 batteries are less than 2000 hours?

Submit Answer

Correct!
Now let's solve the problem after we converted. What is the distribution of the sample mean?
Comment on this question

Select one:

- A t-distribution

- B unknown
- C normal distribution
- D binomial distribution

Submit Answer
THE DISTRIBUTION OF SAMPLE MEAN IS NORMAL BECAUSE THE ORIGINAL DISTRIBUTION OF ONE BATTERY IS NORMAL.

Correct!
Now, let's review how to compute normal probability. What is the probability that a normal random variable with mean 100 and standard devation 25 to be greater than 125? You need to use Z-transformation and use R command pnorm().
Comment on this question

Select one:

- A. 0.14
- B 0.15
- C 0.17
- D 0.16

Submit Answer
P(X>125) = P(Z > (125-100)/25) = 1- P(Z < (125-100)/25) =0.1587 We can compute this in R by typing in 1- pnorm((125-100)/25)

Correct!
As the mean life time of each battery is 800 hours, what's the mean of sample mean of randomly chosen 5 batteries?
Comment on this question

Select one:

- A 800
- B 900
- C. 1000
- D 4000

Submit Answer
The mean of the sample mean is identical to the mean of each random variable. Here, each random variable (life of each battery) has mean 800 hrs.

Correct!
Each of x1, x2, x3,...,xn has standard deviation δ. Then, what's the standard deviation of the sample mean?
Comment on this question

Select one:

- A δ/sqrt(n)

- B square δ
- C δ
- D δ*n

Submit Answer
If each of x1, x2, x3,...,xn has standard deviation δ, then, the standard deviation of the sample mean is δ/sqrt(n)

Correct!
As the life of each battery has standard deviation of 150 hours, what is the standard deviation of the sample mean of 5 batteries?
Comment on this question

Select one:

- A 67.99
- B 67.08
- C 66.99
- D 66.08

Submit Answer
we just apply the property in the previous problem and plug in the number into the formula: δ/sqrt(n). Here, δ=150 hours, n=4

Correct!

Now let's come back to the original problem.

The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the battery life has a normal distribution. Suppose there are 5 batteries on hand. What is the probability that the 5 batteries are used up in less than 4000 hours?

Comment on this question

Select one:

- A 0.975
- B 0.95
- C 0.5
- D unkown

Submit Answer
We know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 67.08. So the problem turns into regular normal probability computation. P(total hours < 4000) = P

(sample mean < 8000) = P(Z < (800-800)/67.08) This can be solved in R by the following command: pnorm((800-800)/67.08)

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24899

You are previewing content.
The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the distribution of battery life is unknown. What is the probability that the 5 batteries are used up in less than 4000 hours?

Comment on this question
Request Help

Select one:

- ○ A 0.95
- ○ B 0.975
- ○ C 0.5
- ◉ D unknown

Submit Answer

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24900

You are previewing content.
The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the distribution of battery life is unknown. Suppose there are 30 batteries on hand. What is the probability that the 30 batteries are used up in less than 25000 hours?

Comment on this question
Request Help

Select one:

- ◉ A 0.87
- ○ B 0.88
- ○ C 0.89
- ○ D 0.90

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!

"What is the probability that the 30 batteries are used up in less than 25000 hours?"

Can you restate this problem in terms of sample mean of battery life, not of the total battery life?

Comment on this question

Select one:

- ○ A. What is the probability that the mean life of 30 batteries are less than 25000 hours?
- ◉ B. What is the probability that the mean life of 30 batteries are less than 853.33 hours?
- ○ C. What is the probability that the mean life of 30 batteries are less than 25000/sqrt(30) hours?

Submit Answer

Correct!
Now let's solve the problem after we converted. What is the distribution of the sample mean?
Comment on this question

Select one:

- A t-distribution
- B normal distribution
- C binomial distribution
- D unknown

Submit Answer

Correct!
Now, let's review how to compute normal probability. What is the probability that a normal random variable with mean 100 and standard devation 25 to be greater than 125? You need to use Z-transformation and use R command pnorm().
Comment on this question

Select one:

- A. 0.14
- B 0.15
- C 0.17
- D 0.16

Submit Answer
P(X>125) = P(Z > (125-100)/25) = 1- P(Z < (125-100)/25) =0.1587 We can compute this in R by typing in 1- pnorm((125-100)/25)

Correct!
As the mean life time of each battery is 800 hours, what's the mean of sample mean of randomly chosen 30 batteries?
Comment on this question

Select one:

- A 950
- B 900
- C 850
- D 800

Submit Answer
The mean of the sample mean is identical to the mean of each random variable. Here, each random variable (life of each battery) has mean 800 hrs.

Correct!
Each of x1, x2, x3,...,xn has standard deviation δ. Then, what's the standard deviation of the sample mean?
Comment on this question

Select one:

- A δ/sqrt(n)

- B δ squared
- C δ
- D δ*n

Submit Answer
If each of x1, x2, x3,...,xn has standard deviation δ, then, the standard deviation of the sample mean is δ/sqrt(n)

Correct!
As the life of each battery has standard deviation of 150 hours, what is the standard deviation of the sample mean of 30 batteries?
Comment on this question

Select one:

- A 27.19
- B 27.29
- C 27.39
- D 27.49

Submit Answer

Correct!

Now, let's come back to the original problem.

The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the distribution of battery life is unknown. Suppose there are 30 batteries on hand. What is the probability that the 30 batteries are used up in less than 25000 hours?

Comment on this question

Select one:

- A 0.871
- B 0.867
- C 0.888
- D 0.901

Submit Answer
We know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 27.39. So the problem turns into regular normal probability computation. P(total hours < 25000) = P (sample mean < 833.33) = P(Z < (833.33-800)/27.39) This can be solved in R by the following command: pnorm((833.33-800)/27.39)

Appendix A

Correct!
You are done with this Assignment!
Comment on this problem

Assistment

Assistment #25996

You are previewing content.
A school district decided that the number of students attending their high school was nearly unmanageable, so they decided to split into two districts, with District 1 students going to the old high school and District 2 students going to a newly constructed building. A group of parents became concerned with how the two districts were constructed relative to income levels. A study was thus conducted to determine whether persons in District 1 have a different mean income from those in District 2. A random sample of 20 homeowners was taken in District 1. Although 20 homeowners were to be interviewed in District 2 also, one person refused to provide the information requested, even though the researcher promised to keep the interview confidential. Thus, only 19 observations were obtained from District 2. The data, recorded in thousands of dollars, produced sample means and variances as shown below. Use these data to construct a 95% confidence interval for $\mu 1-\mu 2$, to two decimal places.

	District 1	District 2
Sample Size	20	19
Sample Mean	18.27	16.78
Sample Variance	8.74	6.58

Comment on this question
Request Help

Type your answer below:

- 9

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
What is the point estimate of this problem?
Comment on this question
Request Help

Select one:

- ○ $\mu 1- \mu 2$ (μ= population mean)
- ○ $\sigma 1- \sigma 2$ (σ= standard deviation)
- ◉ $\hat{Y}$ 1 - $\hat{Y}$ 2 ($\hat{Y}$ = sample mean)
- ○ n1 - n2 (n=sample size)

Submit Answer
No, sorry

Correct!

How many degrees of freedom does the T-distribution of the difference of the means have?

Comment on this question
Request Help

Select one:

- 37
- 28

Submit Answer

Correct!
What is the critical value, to two decimal places?
Comment on this question
Request Help

Type your answer below (mathematical expression):

- 0.004

Submit Answer
No, sorry

Assistment

Assistment #24901

You are previewing content.
Suppose that a tire factory wants to set a mileage guarantee on its new model called LA 50 tire. Life tests indicated that the mean mileage is 47,900, and standard deviation of the normally distributed distribution of mileage is 2,050 miles. The factory wants to set the guaranteed mileage so that no more than 5% of the tires will have to be replaced. What guaranteed mileage should the factory announce?

Comment on this question

Request Help

Select one:

- A 44508
- B 44518
- C 44528
- D 44538
- E 44548

Submit Answer

Correct!
You are done with this problem!

Comment on this problem

Go to next problem

Assistment

Assistment #24902

You are previewing content.
Suppose that the management of a restaurant claimed that 70% of their customers returned for another meal. In a week in which 80 new (first-time) customers dined at the restaurant, what is the probability that 60 or more of the customers will return for another meal?
Comment on this question
Request Help

Select one:

- A 0.1
- B 0.2
- C 0.3
- D 0.4
- E 0.5

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
The answer you types was wrong and we will come back to this topic later in the period. Click ok to continue.
Comment on this question

Select one:

- ok

Submit Answer

Correct!
You are done with this problem!
Comment on this problem
Go to next problem

Assistment

Assistment #24903

You are previewing content.
Suppose you must establish regulations concerning the maximum number of people who can occupy a lift. You know that the total weight of 8 people chosen at random follows a normal distribution with a mean of 550kg and a standard deviation of 150kg. What's the probability that the total weight of 8 people exceeds 600kg?
Comment on this question
Request Help

Select one:

- A 0.37
- B 0.38
- C 0.39
- D 0.40
- E 0.41

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
The answer you types was wrong and we will come back to this topic later in the period. Click ok to continue.
Comment on this question

Select one:

- ok

Submit Answer

Correct!
You are done with this Assignment!
Comment on this problem

Assistment

Assistment #24873

You are previewing content.
For the population of females between the ages of 3 and 74 who participated in the National Health Interview Survey, the distribution of hemoglobin levels is approximately normal with mean 13.3 g/100ml and standard deviation 1.12 g/100ml. If repeated samples are of size 30, what's the probability of these 30 females' hemoglobin levels will have a mean between 13.0 and 13.6 g/100ml?

Comment on this question

Request Help

Select one:

- ○A. 0.975
- ○B. 925
- ○C. 0.745
- ○D. 0.857
- ○E. Cannot be determined

Submit Answer

Assistment

Assistment #24870

You are previewing content.
The weight of anodized reciprocating pistons produced by Brown Company follows a normal distribution with mean 10 lb and standard deviation 0.2 lb. Suppose Brown Company can sell only those pistons weighing between 9.8 and 10.4 lb. What proportion of the pistons is lost?

Comment on this question

Request Help

Select one:

- ○A. 0.75
- ○B. 0.68
- ○C. 0.18
- ○D.0.89
- ○E. 0.92

Submit Answer

Assistment

Assistment #24871

You are previewing content.
Let Y represent the number of years that a chosen undergraduate student in WPI spent obtaining his or her bachelor degree. The mean of Y, which is the mean number of years that students in WPI spent graduating, is 3.8 years, and the variance is 1.44 years squared. One of the statistic professors randomly selects 30 graduated students in WPI from all kinds of major. What's the probability that the total number of years these students spent would be at least 120?

Comment on this question

Request Help

Select one:

- ○A. 0.35
- ○B. 0.23
- ○C. 0.1
- ○D. 0.18
- ○0.57

Submit Answer

Assistment

Assistment #24872

You are previewing content.
In the Netherlands, healthy males between the ages of 65 and 79 have a distribution of serum uric acid levels with meanμ=341 umol/l and standard deviationδ=79 umol/l. What proportion of samples of size 5 have a mean serum uric acid level between 300 and 400 umol/l?

Comment on this question

Request Help

Select one:

- ○ A. 0.95
- ○ B. 0.975
- ◉ C. 0.850
- ○ D. 0.50
- ○ E. Cannot be determined

Submit Answer

Assistment

Assistment #24906

You are previewing content.
A wizard told a prophecy: when a magic light bulb sustain less than 365 days, the couple who used the light bulb will break up within one year. An average magic light bulb lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that a magic light bulb will last at most 365 days?

Comment on this question

Request Help

Select one:

- ○A 0.95
- ○B 0.975
- ○C 0.90
- ○D 0.85
- ○E cannot be decided

Submit Answer

Assistment

Assistment #24905

You are previewing content.
Suppose scores on an IQ test for all the students in a random Calculus 3 class are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?

Comment on this question

Request Help

Select one:

- ○A 0.71
- ○B 0.70
- ○C 0.69
- ○D 0.68
- ○E 0.67

Submit Answer

Assistment

Assistment #25944

You are previewing content.

Have you noticed about the year on your pennies? Sometimes, government would ask bank to recollect pennies on a particular year, say 1957 for some special reason. One penny made in Year 1957 would worth 100 dollar if you choose to give it back to bank. The fewer pennies there are, the more valuable they would be. Are you wondering how many pennies made in 1957 there are and what's the probability you could get one? Click on the website link and you will see the distribution of the number of pennies made in Year 1950 to 2000. http://statweb.calpoly.edu/chance/applets/SampleData/SampleData.html

If type in a positive integer in sample size and number of samples and click on "Draw samples", it would come out the distribution or histogram of number of pennies versus year. If the number you typed in is too large, then try some small ones such as 10 or 20. After you get tired of the process of drawing one penny each time, you could cancel "Animate".

What's the distribution of the number of pennies made in Year 1950 to 2000?

Comment on this question

Request Help

Select one:

- ○ A. normal
- ○ B. skew left
- ○ C. binomial
- ○ D. skew right
- ○ E. unknown

Submit Answer

Assistment

Assistment #25945

You are previewing content.
Suppose you own all the pennies from 1950 to 2000. You pull out 5 pennies each time from them and calculate the mean of their years made, after doing it for 50 times, what's the distribution of mean of the years of those pennies made?

Comment on this question

Request Help

Select one:

- ◯ A. normal
- ◯ B. skew left
- ◯ C. binomial
- ◯ D. skew right
- ◯ E. unknown

Submit Answer

Assistment

Assistment #25946

You are previewing content.
This time you pull out 30 pennies each time and calculate the mean of their years made, after doing it 500 times, (you could ask help to do this experiment from others), what's the distribution of mean of the years of those pennies made?

Comment on this question

Request Help

Select one:

- ◯ A. normal
- ◯ B. skew left
- ◯ C. binomial
- ◯ D. skew right
- ◯ E. unknown

Submit Answer

Assistment

Assistment #25947

You are previewing content.
Again, you are trying to pull out 30 pennies each time and calculate the mean of their years made, after doing it 500 times, what's the probability that the average year of those pennies made would less than Year 1987 when you were born? (express your answer to the nearest 100^{th})

Comment on this question

Request Help

Type your answer below:

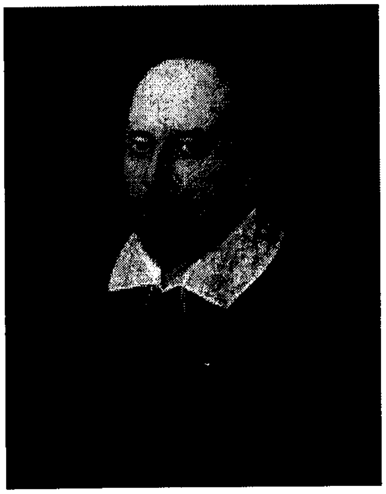
-

Submit Answer

Assistment

Assistment #25950

You are previewing content.



It was told that novels with the same author would contain some words with high frequency, typical length of words and the similar number of nouns. Shakespeare was believed to be a very special writer. Is this also true in his most famous love story - *Romeo and Juliet*? By studying of all the words in *Romeo and Juliet*, the distributions of letters in all the words and proportion of long words as well as proportion of nouns are shown on the link below.

http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

You could try to type in a positive integer in sample size and # of samples. If you click on "Draw samples", it would come out three distributions or histograms: 1) Average length of the sampling words 2) The proportion of long words in the sample 3) the proportion of nouns. If the number you typed in is too large, then try some small ones such as 10 or 20. After you get tired of the process of drawing one word each time, you could cancel "Animate".

What's the distribution of the number of letters in *Romeo and Juliet*?

Comment on this question

Request Help

Select one:

- ○ A. normal
- ○ B. skew left
- ○ C. binomial
- ○ D. skew right
- ○ E. unknown

Submit Answer

Assistment

Assistment #25951

You are previewing content.
If you turn on any page in *Romeo and Juliet* and pick up any 3 words each time and calculate their average length, after doing it 50 times, what's the distribution of the mean number of letters in those words?

Comment on this question

Request Help

Select one:

- ○ A. normal
- ○ B. skew left
- ○ C. binomial
- ○ D. skew right
- ○ E. unknown

Submit Answer

Assistment

Assistment #25953

You are previewing content.
If you turn on any page in *Romeo and Juliet* and this time pick up any 30 words and calculate their average length, after doing it 500 times, what's the probability that the average length of those 30 words would between 4.0 and 4.4 letters? (express your answer to the nearest 1000^{th})

Comment on this question

Request Help

Type your answer below:

-

Submit Answer

Assistment

Assistment #25959

You are previewing content.

http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html Click on the website link and you will see the population distribution of letters in all the words and proportion of long words as well as proportion of nouns. You could try to type in a positive integer in sample size and number of samples. If you click on "Draw samples", it would come out three distributions or histograms: 1) Average length of the sampling words 2) The proportion of long words in the sample 3) the proportion of nouns. If the number you typed in is too large, then try some small ones such as 10 or 20. After you get tired of the process of drawing one word each time, you could cancel "Animate". If the number is too large, the distribution cannot show up.

What's the distribution of the number of nouns in 5 words?

Comment on this question

Select one:

- ○ A. normal
- ○ B. skew left
- ○ C. binomial
- ○ D. skew right
- ○ E. unknown

Submit Answer

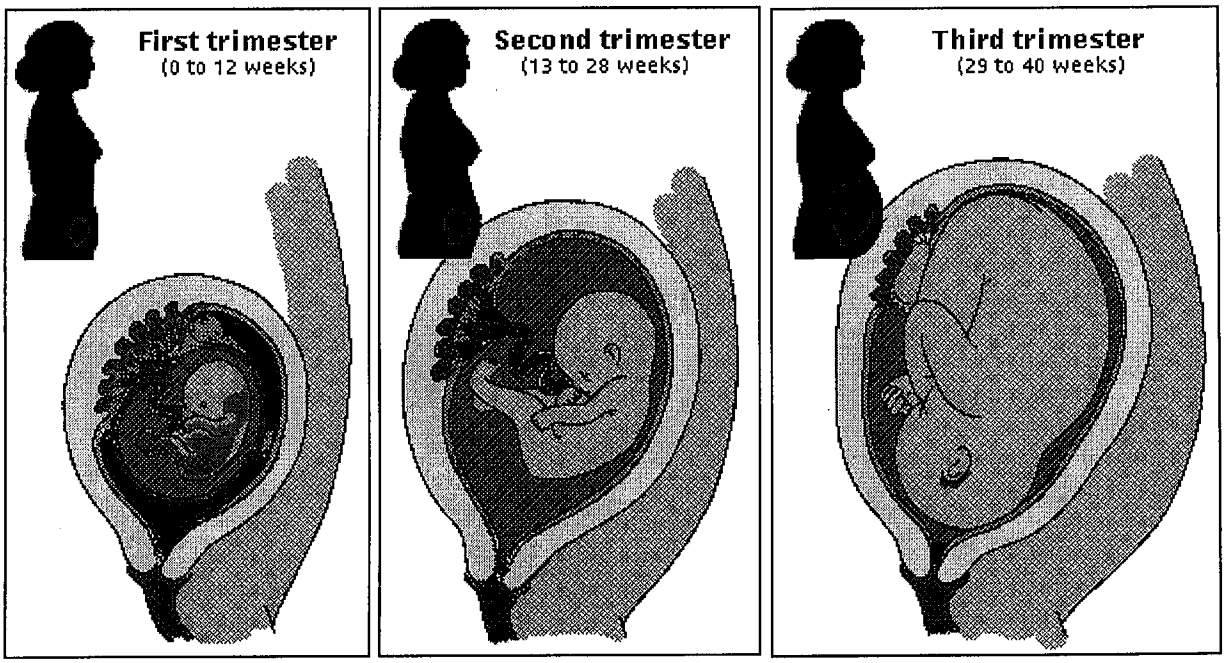
Appendix A

Assistment

Assistment #25774

You are previewing content.

In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is normal, what is the probability that the mean weights of 4 random infants to be greater than 3000 grams? (express your answer to the nearest 100th)



Comment on this question

Request Help

Type your answer below:

- 0.1

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
What kind of distribution does the sample mean follow?

Comment on this question

Request Help

Select one:

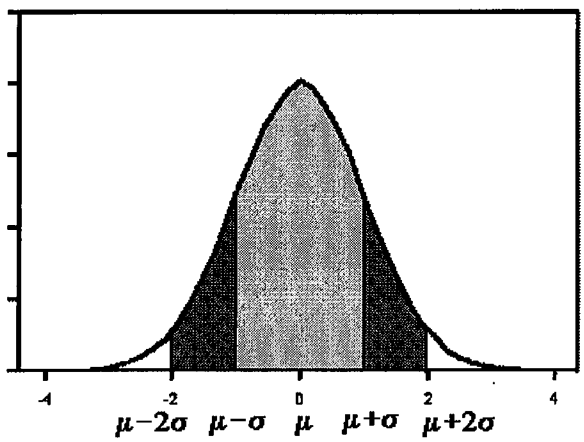
- A. normal distribution

- B. unknown
- C. t-distribution
- D. binomial distribution

Submit Answer
THE DISTRIBUTION OF SAMPLE MEAN WEIGHT IS NORMAL BECAUSE THE DISTRIBUTION OF EACH INFANT'S WEIGHT IS NORMAL.

Correct!
Now, let's review how to compute normal probability. What is the probability that a normal random variable with mean 100 and standard deviation 25 will be greater than 125? You need to use Z-transformation and use R command pnorm(). (Rounding to 2 decimals)
Comment on this question

P(X>125) = P[(X-mean of X)/standard deviation of X > (125- mean of X)/ standard deviation of X]= P [Z > (125-100)/25] = 1- P[Z <(125-100)/25]0.1587. We can compute this in R by typing in 1- pnorm ((125-100)/25)



Comment on this hint

Type your answer below:

- 0.16

Submit Answer
No, sorry

Correct!
We already know how to get a probability from a standard normal distribution. The problem is now to standardize the normal distribution after computing the mean and standard deviation. Thus, what's the mean weight of these 4 babies?
Comment on this question
Request Help

Type your answer below:

Appendix A

- 3500

Submit Answer
No, sorry

Correct!
What's the standard deviation of these 4 infants' mean weights?
Comment on this question

The standard deviation of a sum is the square root of the sum of the squares of the standard deviations:
$SD\,[(X_1+X_2+\ldots+X_n)] = \sqrt{(SD[X1]^2 + SD[X2]^2 + SD[X3]^2 + \ldots + SD[Xn]^2)}$

And we could see n as a constant here, therefore, SD $[\frac{\sum_{i=1}^{n} x_i}{n}]$

$=\sqrt{(\sigma^2 + \sigma^2 + \cdots \sigma^2)/n^2}=\sqrt{n * \sigma^2/n^2}=\sqrt{\sigma^2/n}.$

Here we got n=4, $SD[X_i]$=430.

Comment on this hint

Type your answer below:

- 215

Submit Answer

Correct!
Now let's go back to the original question, "What is the probability that the mean weights of 4 random infants to be greater than 3400 grams?" (Round to 2 decimals)
Comment on this question
We know that the sample mean has 1) normal distribution, 2) mean 3500, and 3) standard deviation 215. So the problem turns into regular normal probability computation. P(sample mean > 3000) = **1-** P (sample mean < 3000) = 1- P(Z < **(3000-3500)/215**) This can be solved in R by the following command: 1-pnorm((3000-3500)/215).
Comment on this hint

Type your answer below:

- 0.99

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25900

You are previewing content.
In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 40 random infants to be greater than 3400 grams?(express your answer to the nearest 1000th)

Comment on this question

Request Help

Type your answer below:

- 0.0

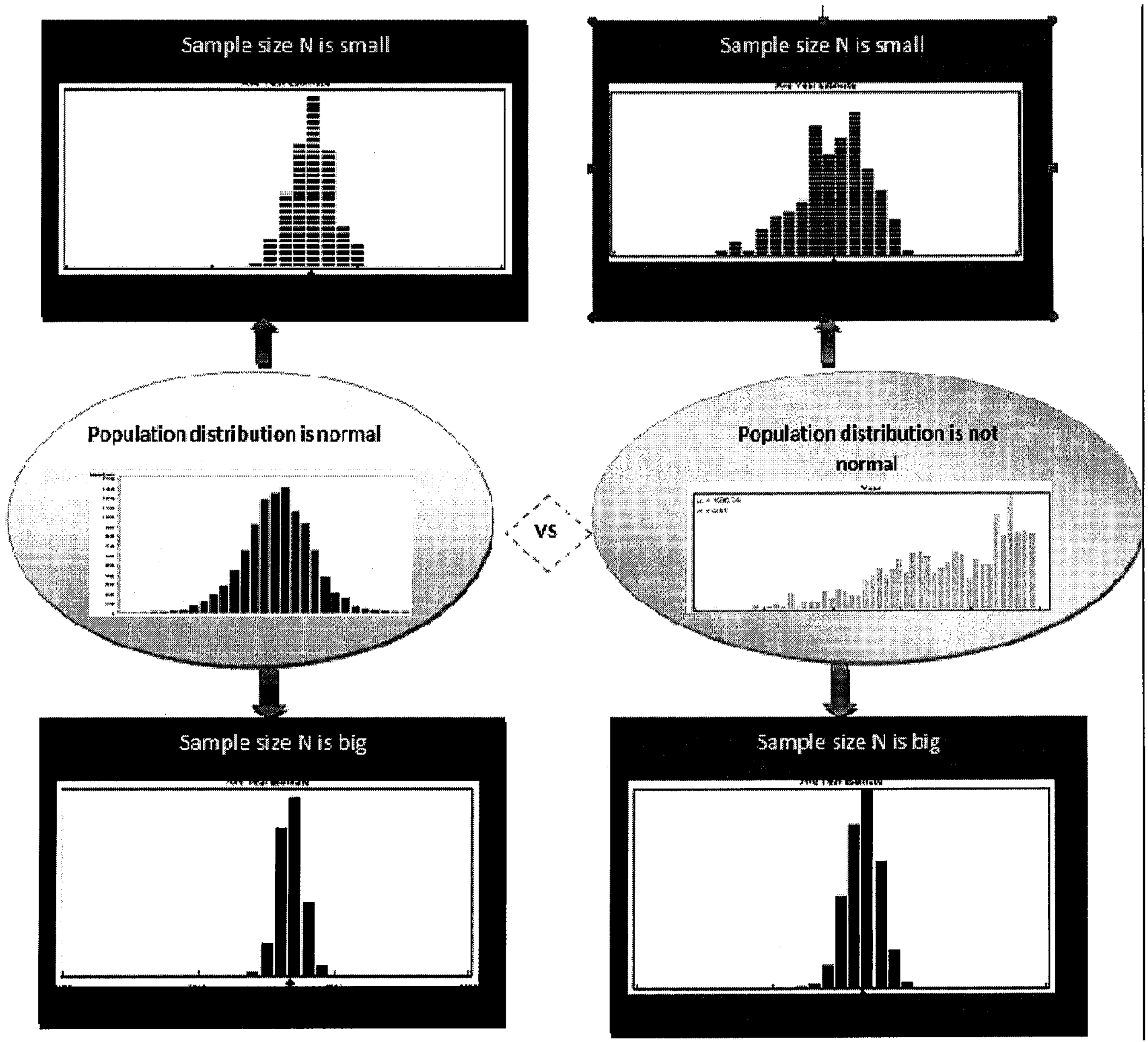
Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!
What is the distribution of the sample mean?

Comment on this question

Even though the question here did not mention what distribution of any infant's birth weight, as the sample size is 40 babies, it is large enough to apply central limit theorem to know the sampling distribution would be normal. The following picture shows this nice theorem's result.

Appendix A



Comment on this hint

Select one:

- A. normal distribution
- B. unknown
- C. t-distribution
- D. binomial distribution

Submit Answer
THE DISTRIBUTION OF SAMPLE MEAN IS NORMAL BECAUSE WE HAVE LARGE ENOUGH SAMPLE SIZE EVEN THOUGH THE ORIGINAL DISTRIBUTION OF ONE INFANT IS UNKNOWN.

Correct!

Appendix A

If we let $B_1, B_2, ..., B_n$ be random variables with mean μ, then what is expected value of $(B_1+B_2+...+B_n)/n$? Since $\mu = 3500$ in our case, what is the expected value of the mean weight of the 40 infants?

A) μ; 3500   B) μ*n; 3500*40   C) $\mu*\sqrt{n}$; 553.4   D) cannot be decided

Comment on this question

The expected value of a sum is the sum of the expected values: $E[\frac{\sum_{i=1}^{n} Bi}{n}] = \sum_{i=1}^{n} E(Bi)/n$

Comment on this hint

Select one:

- A
- B
- C
- D

Submit Answer

That is not correct, try again.

Correct!

Now let $B_1, B_2, ... B_n$ be random variables with standard deviation σ. What is the standard deviation of ?

Since $\sigma = 430$ in this case, what's the standard deviation of the mean weight of the 40 infants?

A) $\sigma/\sqrt{n}$; 67.99   B) $n*\sigma^2$; 7,396,000   C) $n*\sigma$; 17,200   D) $n*\sqrt{\sigma}$; 829.46

Comment on this question

The standard deviation of a sum is the square root of the sum of the squares of the standard deviations:

$SD[(X_1+X_2+...+X_n)] = \sqrt{(SD[X1]^2 + SD[X2]^2 + SD[X3]^2 + ... + SD[Xn]^2)}$

And we could see n as a constant here, therefore, SD $[\frac{\sum_{i=1}^{n} Xi}{n}]$

$= \sqrt{(\sigma^2 + \sigma^2 + \cdots \sigma^2)/n^2} = \sqrt{n * \sigma^2/n^2} = \sqrt{\sigma^2/n}.$

Here, n=40 and $SD[X_i]=430$.

Comment on this hint

Select one:

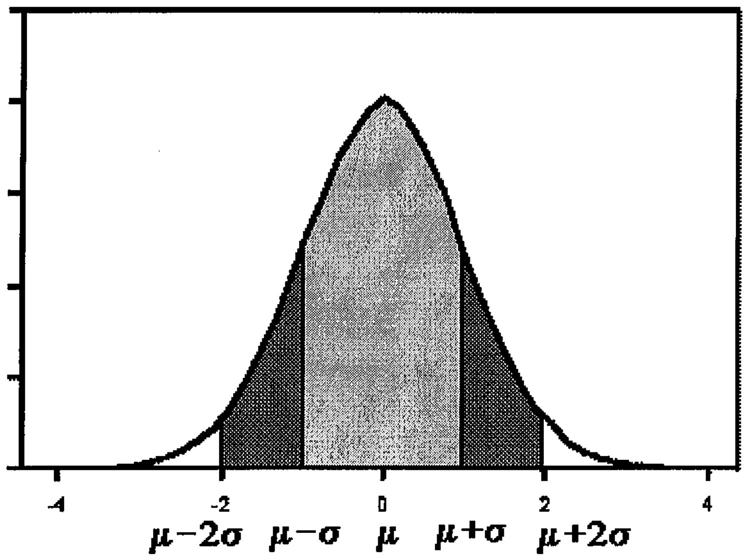
- A
- B
- C
- D

Submit Answer

No, sorry

Correct!
Now let's go back to the original question, "What is the probability that the mean weights of 40 random infants to be greater than 3400 grams?"
Comment on this question

We know that the sample mean has 1) normal distribution, 2) mean 3500, and 3) standard deviation 67.99. So the problem turns into regular normal probability computation. P (sample mean > 3400) = 1- P(sample mean < 3400) = 1- P(Z < (3400-3500)/67.99). This can be solved in R by the following command: 1-pnorm((3400-3500)/67.99).



Comment on this hint

Type your answer below:

- 0.929

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

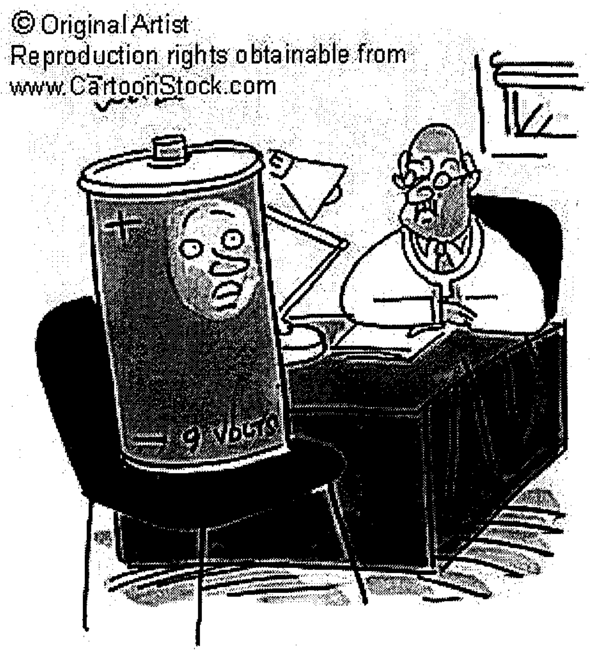
Assistment

Assistment #25913

You are previewing content.

The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the battery life has a normal distribution. Suppose there are 5 batteries on hand. What is the probability that the 5 batteries are used up in less than 3900 hours? (express your answer to the nearest 100th)



"How long have I got left doctor?"

Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!

"What is the probability that the 5 batteries are used up in less than 4000 hours?"

Can you restate this problem in terms of sample mean of battery life, not of the total battery life?

Comment on this question

Select one:

- A. What is the probability that the mean life of 5 batteries is less than 780 hours?
- B. What is the probability that the mean life of 5 batteries is less than 3900 hours?
- C. What is the probability that the mean life of 5 batteries is less than 2000 hours?

Submit Answer
A total hour of 3900 is equivalent to mean of 780 hours.

Correct!

Now let's solve the problem after we converted. What is the distribution of the sample mean?

A. normal distribution

B. unknown

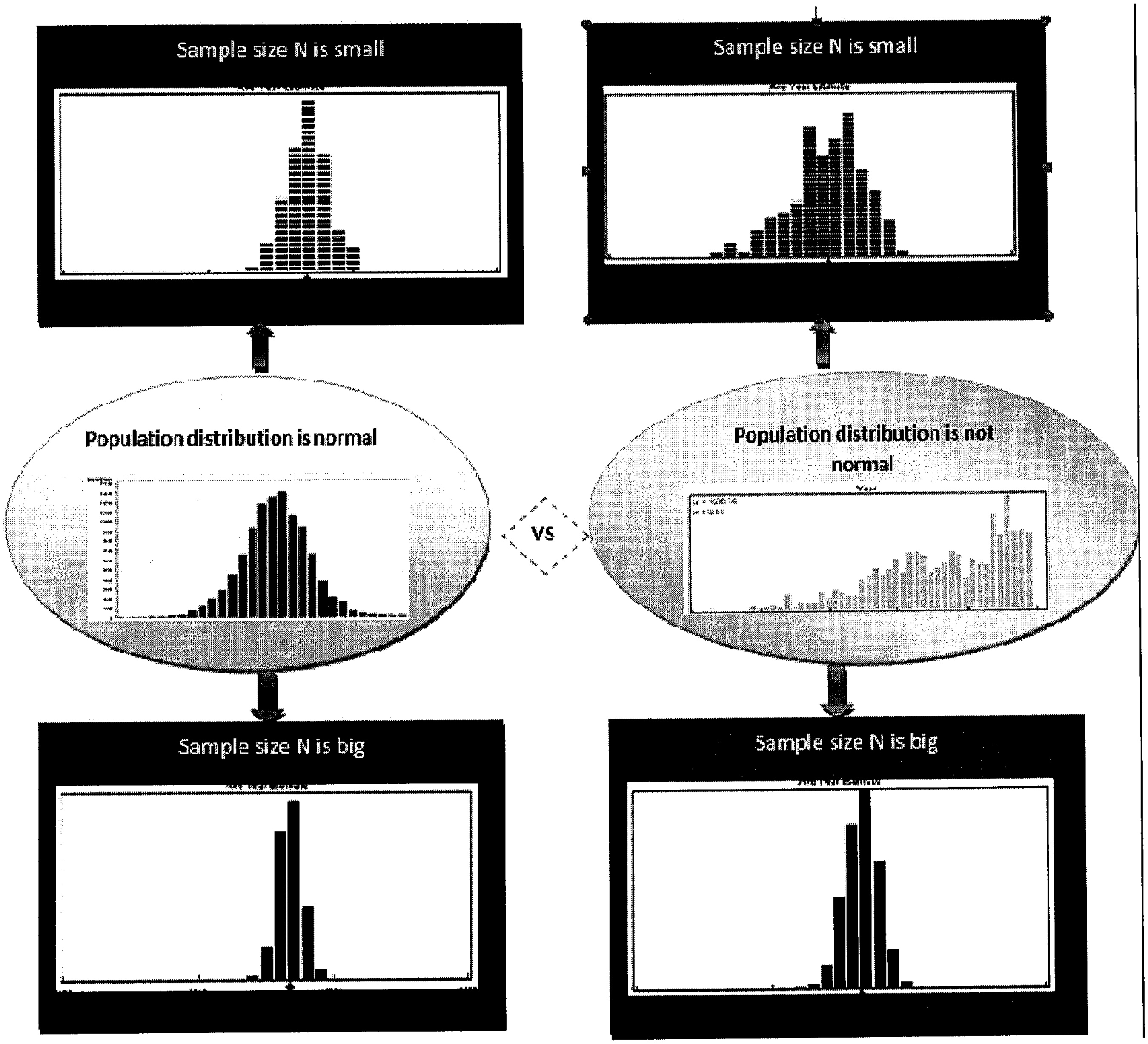
C. t-distribution

D. binomial distribution

Comment on this question

Read the problem again, it said that "assume one battery life's distribution is normal". From the picture below, we can see that no matter how small the sample size is, the sample mean's distribution would be normal as long as the population is normally distributed.

Appendix A



Comment on this hint

Select one:

- A
- B
- C
- D

Submit Answer

Correct!
We know how to get a probability from a standard normal distribution. The problem is now to standardize the normal distribution after figuring out the mean and standard deviation of these 5 batteries' mean lifetime. Thus, what's the mean lifetime of these 5 batteries?

Comment on this question

Appendix A

The expected value of a sum is the sum of the expected values:

$$E[\frac{\sum_{i=1}^{n} Bi}{n}] = \sum_{i=1}^{n} E(Bi)/n$$

Comment on this hint

Type your answer below:

- 800

Submit Answer
No, sorry

Correct!
What's the standard deviation of these 5 batteries' mean lifetime? (rounding to 2 decimals)
Comment on this question

The standard deviation of a sum is the square root of the sum of the squares of the standard deviations:

$$SD[(X_1+X_2+...+X_n)] = \sqrt{(SD[X1]^2 + SD[X2]^2 + SD[X3]^2 + ... + SD[Xn]^2)}$$

And we could see n as a constant here, therefore, $SD\ [\frac{\sum_{i=1}^{n} Xi}{n}]$

$$= \sqrt{(\sigma^2 + \sigma^2 + \cdots \sigma^2)/n^2} = \sqrt{n * \sigma^2/n^2} = \sqrt{\sigma^2/n}.$$

Here, we got n=5, $SD[X_i]=150$.

Comment on this hint

Type your answer below:

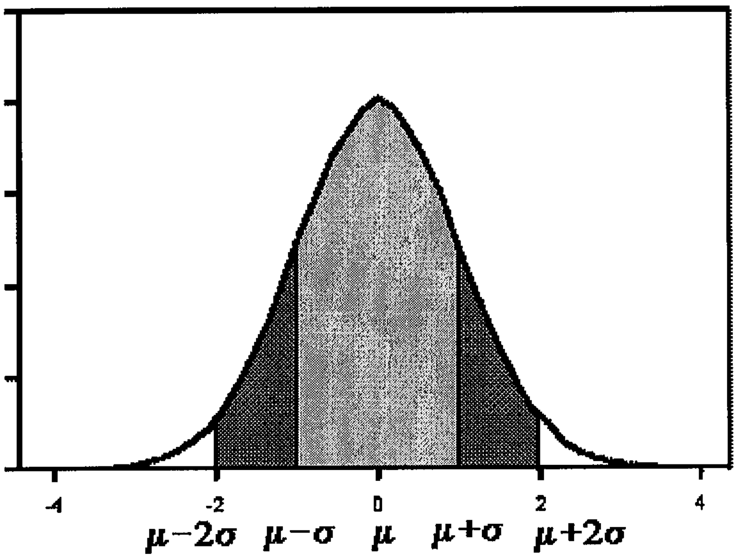
- 67.08

Submit Answer

Correct!
Now let's go back to the original question, "What is the probability that the mean lifetime of 5 random batteries to be less than 780 hours?" (Round to 2 decimals)
Comment on this question

We know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 67.08. So the problem turns into regular normal probability computation. P(total hours < 3900) = P(sample mean < 3900/5) = P(Z < (780-800)/67.08) This can be solved in R by the following command: pnorm((780-800)/67.08).

Appendix A



Comment on this hint

Type your answer below:

- 0.38

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25917

You are previewing content.

The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the distribution of battery life is unknown. What is the probability that the 5 batteries are used up in less than 4000 hours?

Comment on this question

Request Help

Select one:

- ◯ A. 0.95
- ◯ B. 0.5
- ◯ C. 0.975
- ◯ D. cannot be decided

Submit Answer

Let's move on and figure out this problem

Can you restate this problem in terms of sample mean of battery life, not of the total battery life?

A. What is the probability that the mean life of 5 batteries is less than 800 hours?

B. What is the probability that the mean life of 5 batteries is less than 4000 hours?

C. What is the probability that the mean life of 5 batteries is less than $4000/\sqrt{5}$ hours?

Comment on this question

Select one:

- ◉ A
- ◯ B
- ◯ C

Submit Answer

Correct!

Now let's solve the problem after we converted. What is the distribution of the sample mean?

A. normal distribution

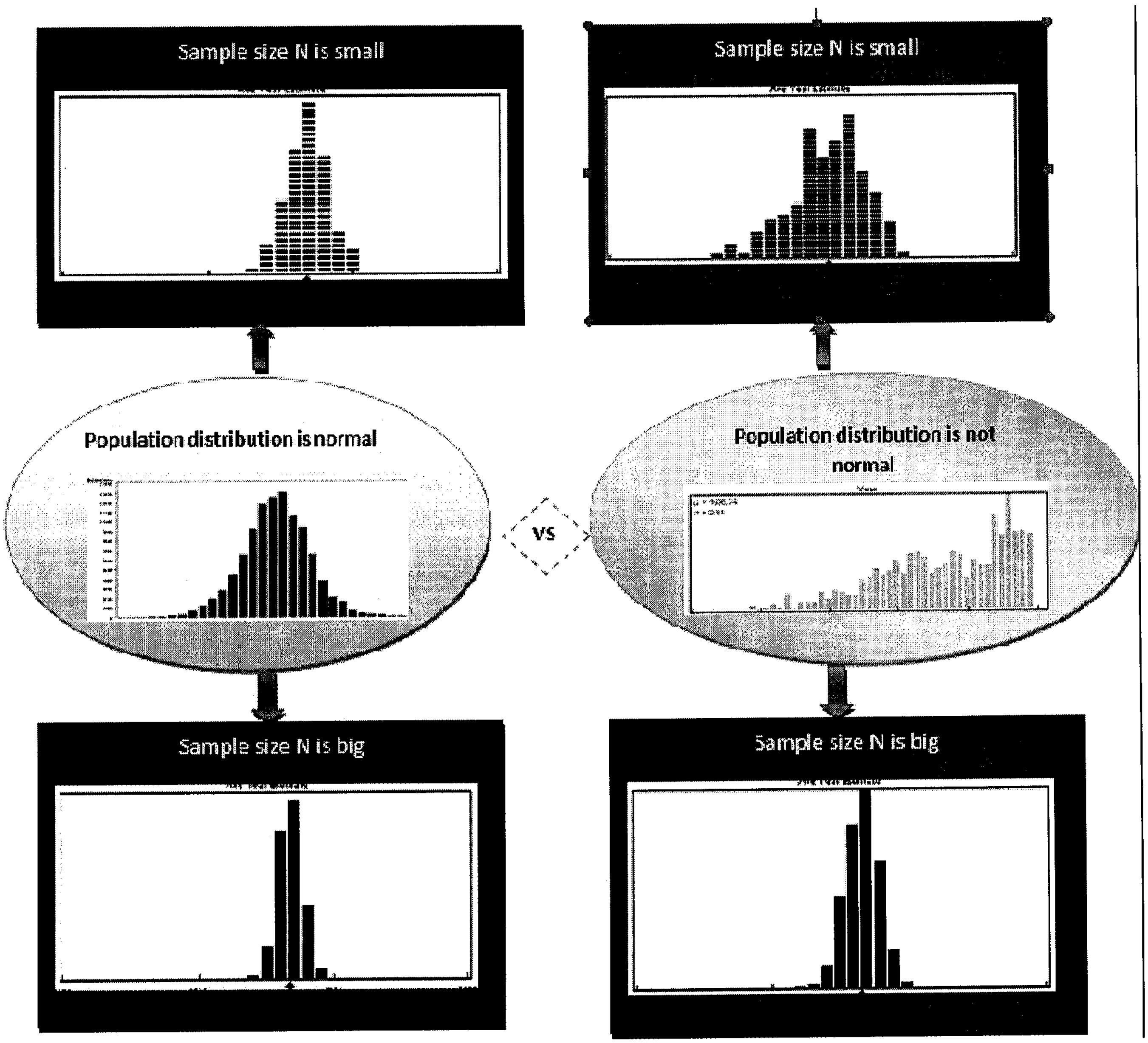
B. unknown

C. t-distribution

D. binomial distribution

Comment on this question

The question assume the distribution of one battery's life time is unknown. From the picture below, we could see that if sample size n=5 is small, the distribution of sample mean is not normal if the population distribution is not normal.



Comment on this hint

Select one:

- ○A
- ◉B
- ○C
- ○D

Submit Answer
THE DISTRIBUTION OF SAMPLE MEAN IS UNKNOWN BECAUSE ORIGINAL DISTRIBUTION OF ONE BATTERY LIFE IS UNKNOWN. UNLESS WE HAVE LARGE ENOUGH SAMPLES TO USE CENTRAL LIMIT THEOREM, WE CANNOT COMPUTE PROBABILITY.

Correct!

What is the probability that the mean life of 5 batteries is less than 4000/5 hours?\A. 0.95

B. 0.5

C. 0.975

D. cannot be decided

Comment on this question

Select one:

- ○A
- ○B
- ○C
- ◉D

Submit Answer

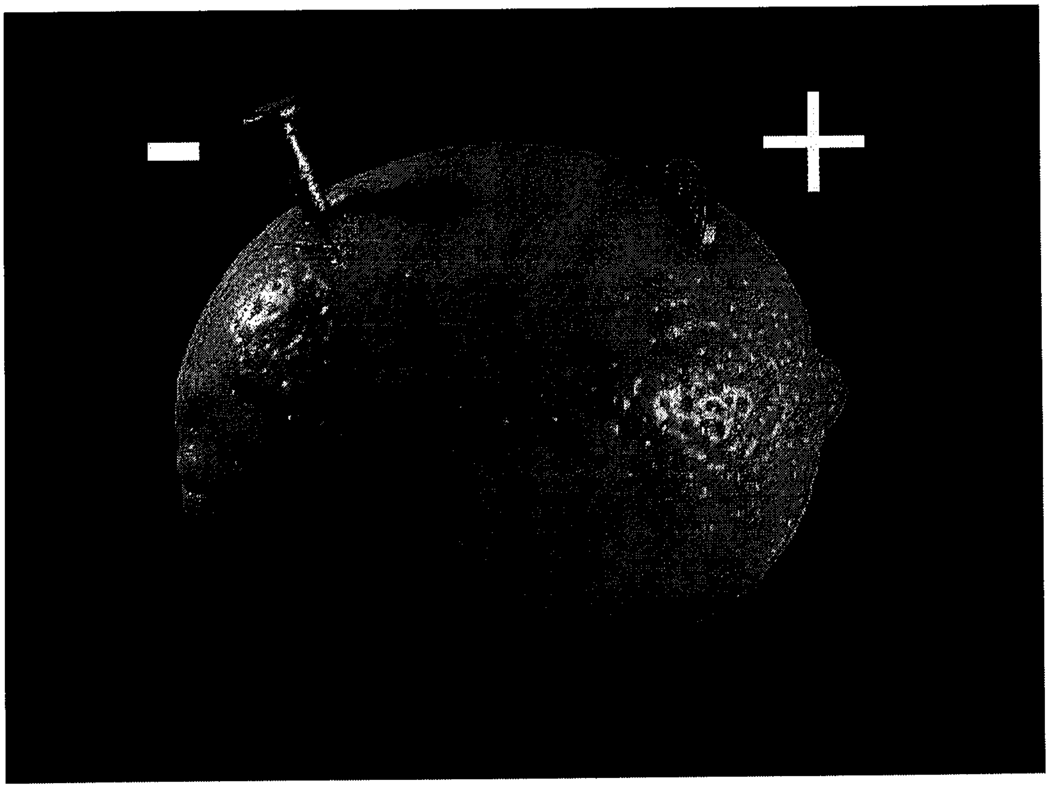
Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25935

You are previewing content.
The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.

Assume that the distribution of battery life is unknown. Suppose there are 30 batteries on hand. What is the probability that the mean life of 30 batteries is less than 853.33 hours?(round to 2 decimal places)



Comment on this question

Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem

What is the distribution of the sample mean?

A. normal distribution

B. unknown

C. t-distribution

D. binomial distribution

Comment on this question

Request Help

Select one:

- ◉ A
- ○ B
- ○ C
- ○ D

Submit Answer

Correct!
We know how to get a probability from a standard normal distribution. The problem is now to standardize the normal distribution after figuring out the mean and standard deviation of these 30 batteries' mean lifetime. Thus, what's the mean lifetime of these 30 batteries?

Comment on this question

The expected value of a sum is the sum of the expected values:

$$E\left[\frac{\sum_{i=1}^{n} Bi}{n}\right] = \sum_{i=1}^{n} E(Bi)/n$$

Comment on this hint

Type your answer below:

- 800

Submit Answer

That is not correct, try again.

Correct!
What's the standard deviation of these 30 batteries' mean lifetime? (Round to 2 decimal)

Comment on this question

The standard deviation of a sum is the square root of the sum of the squares of the standard deviations:

$$SD[(X_1+X_2+...+X_n)] = \sqrt{(SD[X1]^2 + SD[X2]^2 + SD[X3]^2 + ... + SD[Xn]^2)}$$

And we could see n as a constant here, therefore, SD $[\frac{\sum_{i=1}^{n} X_i}{n}]$

$$= \sqrt{(\sigma^2 + \sigma^2 + \cdots \sigma^2)/n^2} = \sqrt{n * \sigma^2/n^2} = \sqrt{\sigma^2/n}.$$

Here, we got n=30, $SD[X_i]=150$.

Comment on this hint

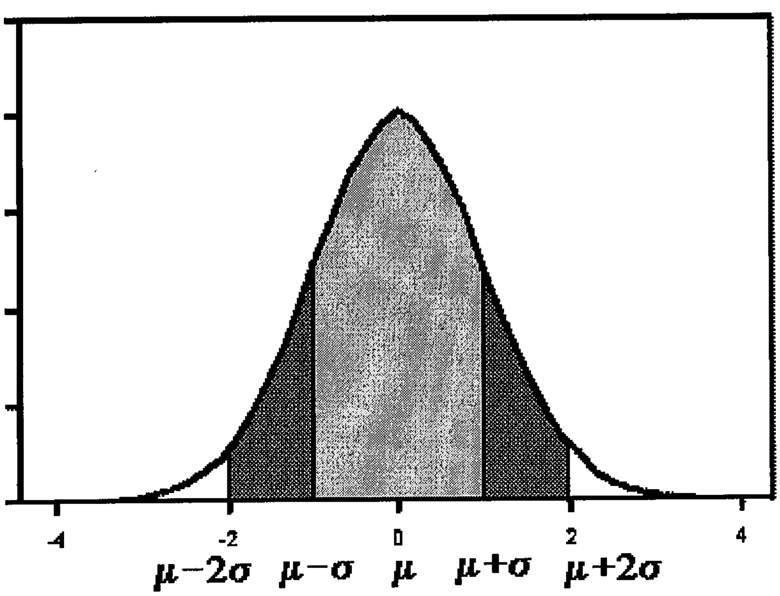
Type your answer below:

- 27.39

Submit Answer

Correct!
Now let's go back to the original question, "What is the probability that the mean lifetime of 30 random batteries to be less than 853.33 hours?" (Round to 2 decimals)
Comment on this question

We know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 27.39. So the problem turns into regular normal probability computation. P(total hours < 25000) = P (sample mean < 833.33) = P(Z < (833.33-800)/27.39) This can be solved in R by the following command: pnorm((833.33-800)/27.39).



Comment on this hint

Type your answer below:

- 0.89

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Appendix A

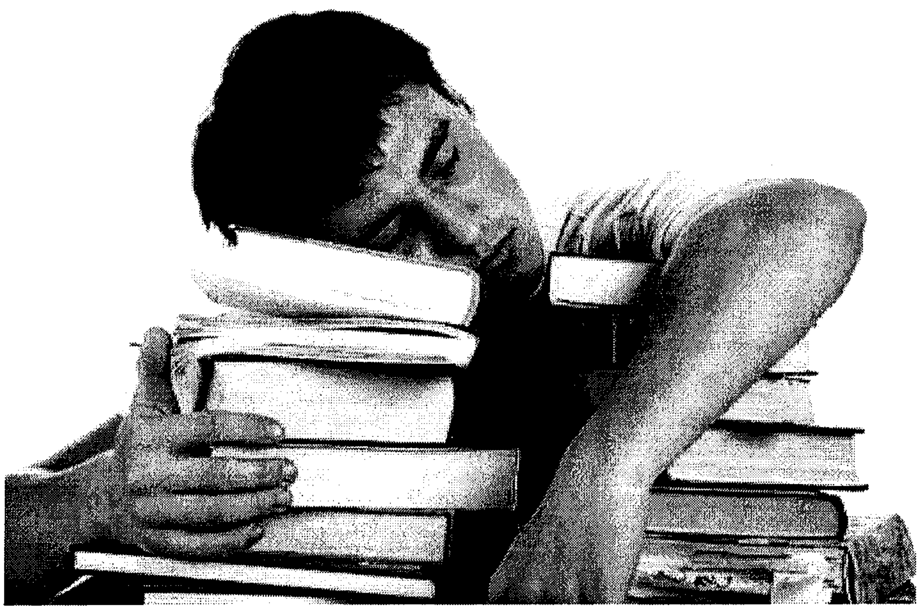
Assistment

Assistment #25936

You are previewing content.

In 2004, the National Survey of Student Engagement surveyed 620,000 students at 850 four-year universities to find out how college students really spent their days. According to the McKinley Health Center at the University of Illinois in Champaign, the college student gets averagely **6** hours of sleep a night and a big **variation** of 16 hours.

To stay rested and at your best, aim to get at least eight hours of sleep. What's the probability that the mean sleeping time of 30 random chosen students would achieve this goal? (rounding to 3 decimals)



Comment on this question

Request Help

Type your answer below:

- 0.01

Submit Answer

Sorry, that is incorrect. Let's move on and figure out why!

What is the distribution of the sample mean?

A. unknown

B. normal distribution

C. t-distribution

D. binomial distribution

Comment on this question
Request Help

Select one:

- A
- B
- C
- D

Submit Answer

Correct!
We know how to get a probability from a standard normal distribution. The problem is now to standardize the normal distribution after figuring out the mean and standard deviation of these 30 batteries' mean lifetime. Thus, what's the mean sleeping hour of these 30 students?
Comment on this question

The expected value of a sum is the sum of the expected values:

$$E\left[\frac{\sum_{i=1}^{n} Bi}{n}\right] = \sum_{i=1}^{n} E(Bi)/n$$

Comment on this hint

Type your answer below:

- 6

Submit Answer

Correct!
What's the standard deviation of these 30 students' mean sleeping hour? (Round to 2 decimal)
Comment on this question

The standard deviation of a sum is the square root of the sum of the squares of the standard deviations:

$$SD\ [(X_1 + X_2 + \ldots + X_n)] = \sqrt{(SD[X1]^2 + SD[X2]^2 + SD[X3]^2 + \ldots + SD[Xn]^2)}$$

And we could see n as a constant here, therefore, SD $\left[\frac{\sum_{i=1}^{n} Xi}{n}\right]$

$$= \sqrt{(\sigma^2 + \sigma^2 + \cdots \sigma^2)/n^2} = \sqrt{n * \sigma^2/n^2} = \sqrt{\sigma^2/n}.$$

Here, we got n=30, and $SD[X_i]=16$

Comment on this hint

Type your answer below:

- 0.73

Submit Answer

Correct!
Now let's go back to the original question, "What's the probability that the mean sleeping time of 30 random chosen students would achieve this goal?" (Round to 3 decimals)
Comment on this question
Request Help

Type your answer below:

- 0.003

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25939

You are previewing content.

Please determine whether the following statements about central limit theorem are true or not.

1)"The theorem says that the sample mean of n obervations $X_1,\ldots X_n$, is approximately normally distributed."

2) "The central limit theorem guarantees that n observations in a random sample will be approximately normally distributed if n is sufficiently large."

(True=T, False=F)

Comment on this question

1) false, because Central limit theorem requires n to be large enough.

2) The theorem guarantees that the **mean** of a random sample follows normal distribution if n is large.

Comment on this hint

Select one:

- ○ A. TT
- ○ B. TF
- ○ C. FT
- ◉ D. FF

Submit Answer
That is not correct, try again.

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25940

You are previewing content.

In most Asian countries, the age-old bias for boy babies has led to the gender imbalance. Combined with the one-child policy, China has the largest, the highest, and the longest gender imbalance in the world. From a relatively normal ratio of 108.5 boys to 100 girls in the early 80s, the male surplus progressively rose to 111 in 1990, 116 in 2000, and is now is close to 120 boys for each 100 girls at the present time, according to a Chinese think-tank report. What's the probability that out of these 1500 single-child families, at least 850 families have boys? (express your answer in the nearest 1000th).

Comment on this question

Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem

What is the distribution of the number of families with boys?

A. unknown

B. t-distribution

C. binomial distribution

Comment on this question

Select one:

- ○ A
- ○ B
- ◉ C

Submit Answer

For each single-child family, it would be a boy or girl baby. This is the same as flipping a coin. Therefore, every trial is a Bernoulli distribution. The sum of Bernoulli distribution is a binomial distribution.

Correct!

Appendix A

We know that Y, the number of families with boys, follows binomial distribution. What's p in this problem? p is the probability for each single-child family to have a boy. (rounding to 2 decimals)

Comment on this question

The probability of having a boy in each family is 120/(120+100).

Comment on this hint

Type your answer below:

- 0.55

Submit Answer

Correct!

What' the mean of Y, the number of families in the sample with boys?

Comment on this question

Y~binomial (n,p). E(Y) =n*p, in this problem, n=1500, p=0.55

Comment on this hint

Type your answer below:

- 825

Submit Answer

Correct!

What' the standard deviation of Y, the number of families in the sample with boys?(express your answer to the nearest 100th).

Comment on this question

Y~binomial (n,p). SD(Y) =√n*p*(1-p),in this problem, n=1500, p=0.55.

Comment on this hint

Type your answer below:

- 19.27

Submit Answer

Correct!

What's the probability that out of these 1500 single-child families, at least 850 families have boys?

Comment on this question

Since n*p=825 >> 5, there is no doubt that we could apply central limit theorem in this case. Therefore, Y approximately follows normal distribution with mean 825 and standard deviation 19.27. So the problem turns into regular normal probability computation. P (Y 900) = P (Z (900-825)/19.27). This can be solved in R by the following command: 1-pnorm ((850-825)/19.27)

Comment on this hint

Type your answer below:

- 0.097

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25941

You are previewing content.
According to government data, 25% of American children under the age of 6 live in households with income less than the official poverty level. A random sample of 300 children under the age of 6 is selected for a study of learning in early childhood. Approximately how likely it is that at least 90 of the children in the sample live in poverty? (express your answer to the nearest $1000^{th)}$

Comment on this question

Request Help

Type your answer below:

-

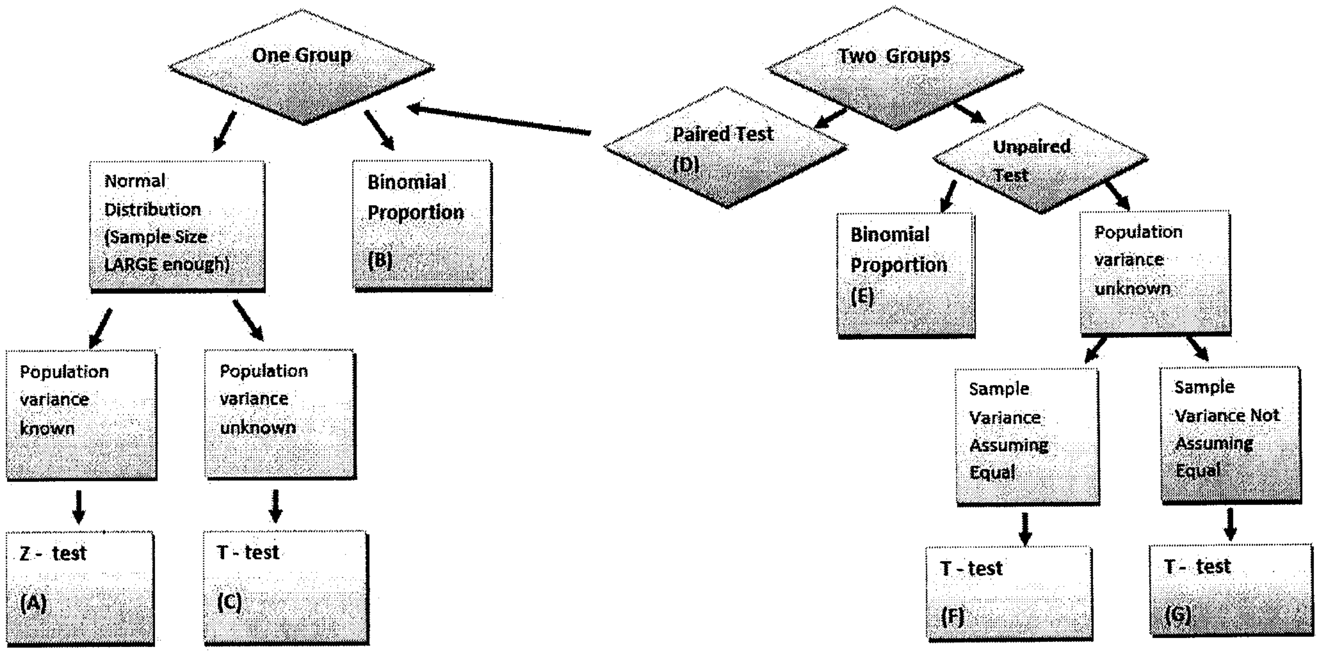
Submit Answer

Assistment

Assistment #25918

You are previewing content.

The mean length of a work week for the population of workers was reported to be 39.2 hours (Investors Business Daily, September 11, 2000). Suppose that we would like to take a current sample of workers to see whether the mean length of a work week has changed from the previously reported 39.2 hours.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question

Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem
The distribution of the population is ...

Comment on this question

Select one:

- Known

- Unknown

Submit Answer
That is not correct, try again.

Correct!
The size of the population is ...
Comment on this question

Select one:

- Large
- Small

Submit Answer

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- Measurement
- Proportion

Submit Answer
No, sorry

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25949

You are previewing content.

Recall the problem above: the mean length of a work week for the population of workers was reported to be 39.2 hours. Suppose a current sample of 112 workers provided a sample mean of 38.5 hours. Use a population standard deviation s = 4.8 hours.
Consider the hypothesis test to see whether the mean length of a work week has changed from the previously reported 39.2 hours. You determined the appropriate hypotheses to be:

H_0: $\mu = 39.2$

H_A: $\mu \neq 39.2$

You determined the p-value for this two sided test to be .1236. Which of the following statements is true?

a) H_0 should be rejected in favor of H_A at the .05 level.

b) H_0 should not be rejected in favor of H_A at the .05 level.

c) There is not enough information to draw either of the above conclusions.

Comment on this question
A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint
Request Help

Select one:

- ○ A
- ◉ B
- ○ C

Submit Answer
That is not correct, try again.

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25937

You are previewing content.

The mean length of a work week for the population of workers was reported to be 39.2 hours (Investors Business Daily, September 11, 2000). Suppose that we would like to take a current sample of workers to see whether the mean length of a work week has changed from the previously reported 39.2 hours.

State the hypotheses that will help us determine whether a change occurred in the mean length of a work week.

A) $H_0: \overline{X} = 39.2$

$H_A: \overline{X} \neq 39.2$

B) $H_0: \mu = 39.2$

$H_A: \mu \neq 39.2$

C) $H_0: \mu \geq 39.2$

$H_A: \mu < 39.2$

D) $H_0: \mu \leq 39.2$

$H_A: \mu > 39.2$

Comment on this question

There are two hypotheses in a hypothesis test: the null hypothesis and the alternative hypothesis. Hypothesis tests are always about population parameters.

Comment on this hint

Request Help

Select one:

- A
- B
- C
- D
-

Submit Answer

Sorry, that's not correct. The problem asks you to test whether the population mean has *changed from* a value. Put another way, we're interested in knowing if the population mean is now different than it was before. The alternative hypothesis here test's to see if the population mean is *smaller* than it was before.

Correct!
You are done with this problem!
Comment on this problem

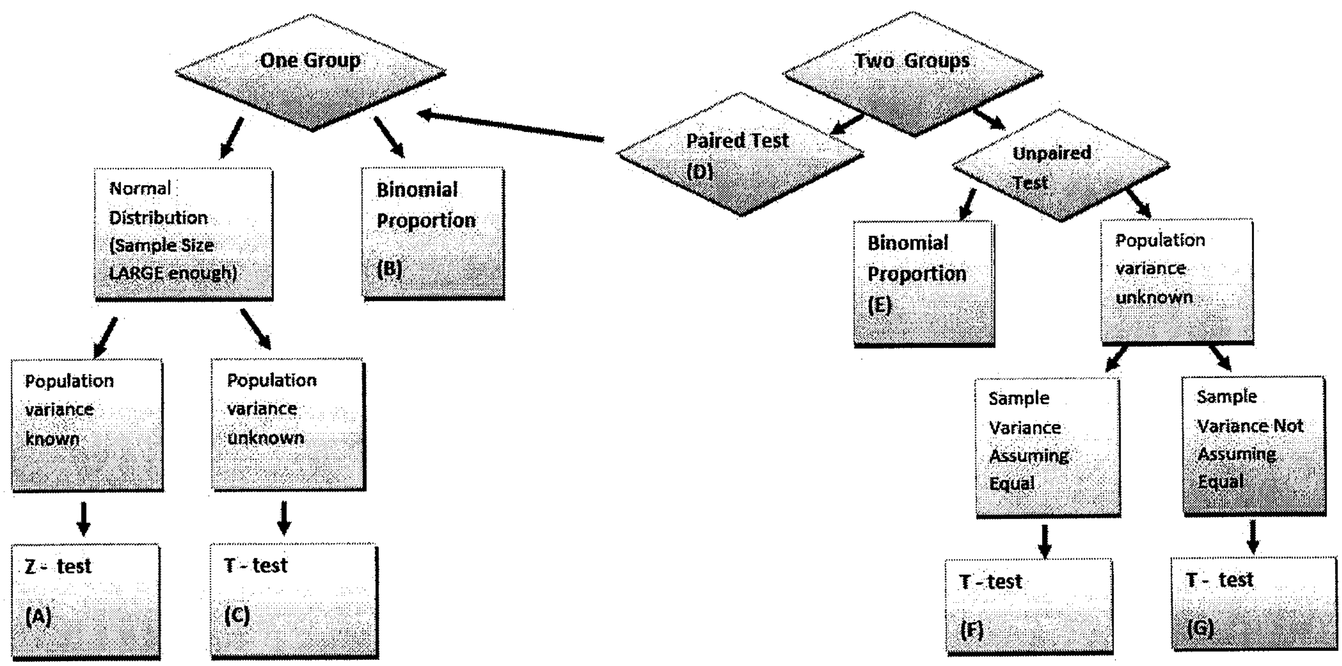
Appendix A

Assistment

Assistment #25955

You are previewing content.

The average Russian woman has 1.17 babies (*A second baby? Russia's mothers aren't persuaded.* Fred Weir. The Christian Science Monitor http://www.csmonitor.com/2006/0519/p01s04-woeu.html). Suppose that we would like to take a current sample of 17 women to see whether the mean number of babies per woman is below the healthy replacement rate of 2.08 babies per woman.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question

Request Help

Select one:

- ◌ A
- ◌ B
- ◌ C
- ◌ D
- ◌ E
- ◌ F
- ◌ G

Submit Answer

Let's move on and figure out this problem
The distribution of the population is ...
Comment on this question

Select one:

- ◌ Known
- ◉ Unknown

Submit Answer
No, try again

Correct!
The size of the population is ...
Comment on this question

Select one:

- ◌ Large
- ◉ Small

Submit Answer
No, try again

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- ◉ Measurement
- ◌ Proportion

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #25961

You are previewing content.

The average Russian woman has 1.17 babies (*A second baby? Russia's mothers aren't persuaded.* Fred Weir. The Christian Science Monitor http://www.csmonitor.com/2006/0519/p01s04-woeu.html). Suppose that we would like to take a current sample of 17 women to see whether the mean number of babies per woman is below the healthy replacement rate of 2.08 babies per woman.

State the hypotheses that will help us determine whether a change occurred in the average number of babies per woman is less than the health replacement rate.

A) H_0: $\overline{X} = 2.08$

H_A: $\overline{X} \neq 2.08$

B) H_0: $\mu = 2.08$

H_A: $\mu \neq 2.08$

C) H_0: $\mu \geq 2.08$

H_A: $\mu < 2.08$

D) H_0: $\mu \leq 2.08$

$H_A: \mu > 2.08$

Comment on this question

There are two hypotheses in a hypothesis test: the null hypothesis and the alternative hypothesis. Hypothesis tests are always about population parameters.

Comment on this hint

Request Help

Select one:

- ○A
- ◉B
- ○C
- ○D

Submit Answer

Sorry, that's not correct. The problem asks you to test whether the population mean is below a value. Put another way, we're interested in knowing if the population mean is now smaller than some number.The alternative hypothesis here test's to see if the population mean is different from a number.

Assistment

Assistment #25962

You are previewing content.
Recall the problem above: the healthy replacement rate of a human population is 2.08 babies per woman. Suppose a sample of 17 women have a total of 20 babies,and the standard deviation of the sample is 0.9396 babies per woman.

Perform a hypothesis test to see whether the mean number of babies per woman is less than the replacement rate of 2.08 babies per woman, and report the p-value for this test below. Please provide your answer to four decimal places.

Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
Which test statistic should we use for this problem: z or t?
Comment on this question
Request Help

Select one:

- ○ z
- ◉ t

Submit Answer
No, sorry

Correct!
Calculate the test or sample value for t (to two decimal places).
Comment on this question

The expression for t* in this case is

$t^* = (\bar{x} - \alpha)/\sigma_{\bar{x}}$, where $\bar{x}$ is the sample mean, α is the value we're testing for and $\sigma_{\bar{x}} = s/\sqrt{n}$ is the standard error (also called the standard deviation of the sampling distribution of the mean.

Comment on this hint

The expression for t in this case is

t* = ($\overline{x}$ -α)/$\sigma_{\overline{x}}$, where $\overline{x}$ is the sample mean (1.176), α is the value we're testing for (2.08) and $\sigma_{\overline{x}}$ = s/√n is the standard error (.9396/√17 = 0.2278).

Comment on this hint

t* = -3.96

Comment on this hint

Type your answer below:

-

Submit Answer

Appendix A

Assistment

Assistment #26023

You are previewing content.

http://www.nytimes.com/2007/01/16/health/nutrition/16brody.html

Does color preference affect weight loss? 41 boys and 63 girls participated in a study in which they lost weight. Of the 69 who preferred the color blue, the average weight loss was 51 pounds, while to overall average weight loss was 59. Assume that the weight loss of the two groups follows $X_{1,i} \sim N(\mu_1, \sigma)$ and $X_{2,i} \sim N(\mu_2, \sigma)$, determine the p-value of color preference.

What type of test should be used to calculate this p-value?

Comment on this question

Request Help

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Assistment

Assistment #26030

You are previewing content.
Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.
Comment on this hint

Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- (selected) There was a difference in weight loss between the group that prefered the color blue and everyone else.
- There was not a difference in weight loss between the group that prefered the color blue and everyone else.
- There is not enough information to draw either of the above conclusions.

Submit Answer
That is not correct, try again.

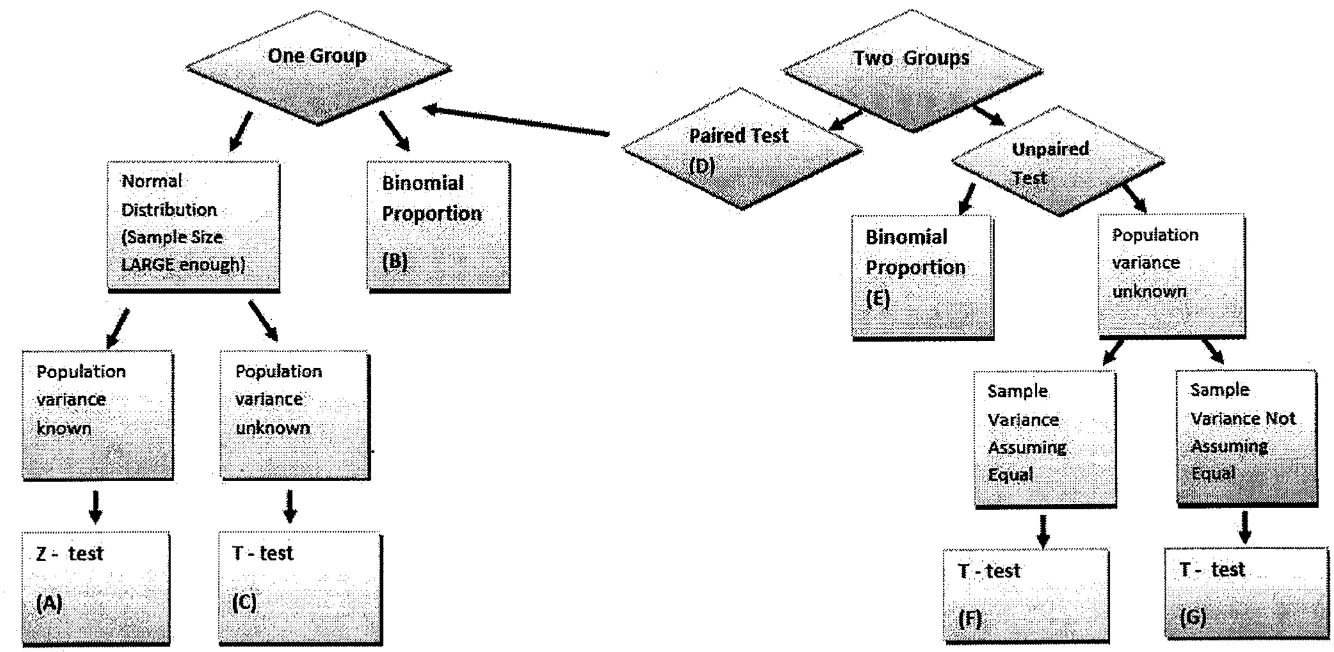
Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26031

You are previewing content.

http://www.journalsleep.org/ViewAbstract.aspx?citationid=3422
If you're a tired college, should you consume caffeine, or sleep? You measure the number of times drowsy drivers cross the lane dividers over two miles. Of the 24 participants, 8 where given decaf, 8 where given 200mg of caffeine in a cup of coffee and 8 where given half an hour to nap. We wish to establish whether the difference between the number of line crossings for the caffeine group versus the nap group is positive.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question
Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem

Appendix A

The number of samples is ...

Comment on this question

Select one:

- One
- Two

Submit Answer
No, sorry

Correct!
Are the individuals in the population ...
Comment on this question

Select one:

- Paired
- Unpaired

Submit Answer
No, sorry

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- Measurement
- Proportion

Submit Answer

Correct!
Are the standard deviations of the two distributions assumed to be the same?
Comment on this question

Select one:

- Yes
- No

Submit Answer
That is not correct, try again.

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26034

You are previewing content.

Recall that we'd like to establish whether the difference between the average number of line crossings for the eight person caffeine group versus the nap group, of the same size, is positive.

Consider the hypothesis test to see whether the difference between the average number of line crossings for the eight person caffeine group versus the nap group, of the same size, is positive.

You determined the appropriate hypotheses to be:

$H_0: \mu_1 \leq \mu_2$

$H_A: \mu_1 > \mu_2$

You determined the p-value for this one sided test to be .0694 Which of the following statements is true?
Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.

Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ○ The difference between the average number of line crossings for the caffeine group versus the nap group is positive.
- ○ The difference between the average number of line crossings for the caffeine group versus the nap group is not positive.
- ◉ There is not enough information to draw either of the above conclusions.

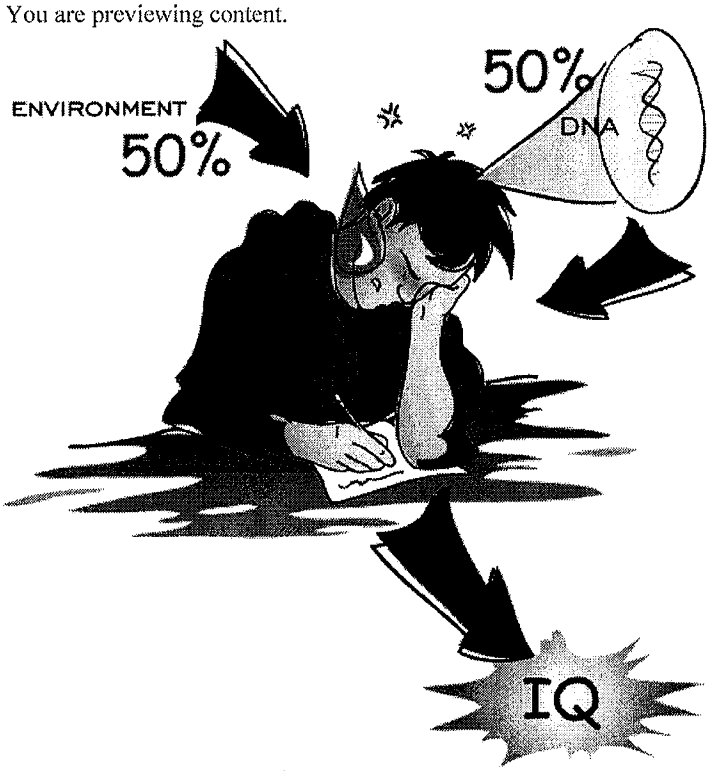
Submit Answer
No, try again

Correct!
You are done with this problem!
Comment on this problem

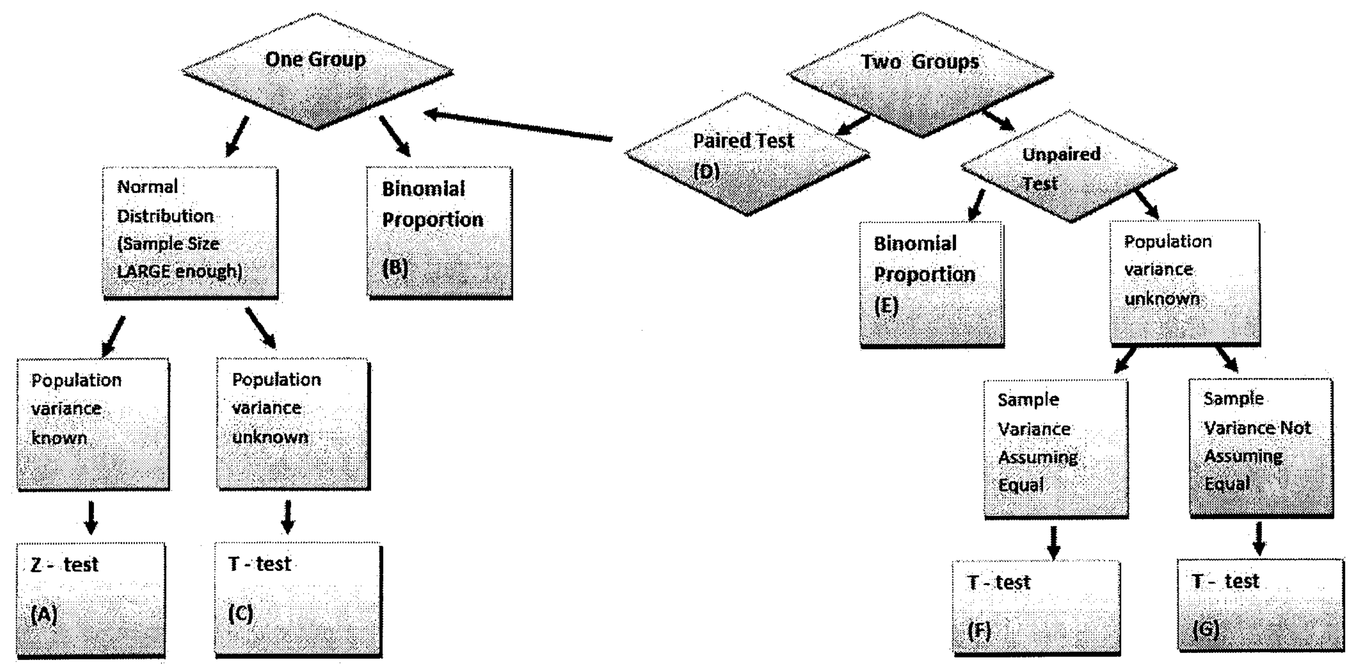
Assistment

Assistment #26035

You are previewing content.



http://www.sciencemag.org/cgi/content/full/316/5832/1717
Are older siblings smarter than their younger counter parts? IQ tests are administered to the children of 716 families with two kids and the intra-family differences are compared. Assume that the standard deviation of the differences will be given.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question
Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem
The distribution of the population is ...
Comment on this question

Select one:

- Known
- Unknown

Submit Answer
No, sorry

Correct!
The size of the population is ...
Comment on this question

Select one:

- Large
- Small

Submit Answer

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- Measurement
- Proportion

Submit Answer
That is not correct, try again.

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26036

You are previewing content.

State the hypotheses that will help us determine whether there is a positive difference between the older siblings and the younger one's IQ scores.

(a) H_0: $\overline{X}_1 = \overline{X}_2$

H_A: $\overline{X}_1 \neq \overline{X}_2$

(b) H_0: $\mu = 0$

H_A: $\mu \neq 0$

(c) H_0: $\mu \geq 0$

H_A: $\mu < 0$

(d) H_0: $\mu \leq 0$

H_A: $\mu > 0$

Comment on this question

There are two hypotheses in a hypothesis test: the null hypothesis and the alternative hypothesis. Hypothesis tests are always about population parameters.

Comment on this hint

We believe the null hypothesis unless there is significant enough evidence to convince us to reject it in favor of the alternative hypothesis.

Comment on this hint

H_0: $\mu \leq 0$

H_A: $\mu > 0$

Comment on this hint

Select one:

- ◯ A
- ◯ B
- ◯ C
- ◉ D

Submit Answer

Sorry, that's not correct. The problem asks you to test whether there is a positive difference in the corresponding IQ scores. Put another way, we're interested in knowing if the population mean of the difference is positive. The alternative hypothesis here test's to see if the population mean is less than zero

Correct!

You are done with this problem!

Comment on this problem

Assistment

Assistment #26038

You are previewing content.
Recall the problem above: we want to determine whether there is a positive difference between the older siblings and the younger one's IQ scores.

We administered an IQ test to the oldest and second oldest child of 716 families. Suppose the average difference is found to be 2 IQ points. with a variance of 2.17

Consider the hypothesis test to see whether there is a positive difference between the older siblings and the younger one's IQ scores. You determined the appropriate hypotheses to be:

H_0: $\mu \leq 0$

H_A: $\mu > 0$

You determined the p-value for this one sided test to be .0873. Which of the following statements is true?
Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.

Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ○ There is a positive difference between the older siblings and the younger one's IQ scores
- ○ There is not a positive difference between the older siblings and the younger one's IQ scores
- ◉ There is not enough information to draw either of the above conclusions.

Submit Answer
No, try again

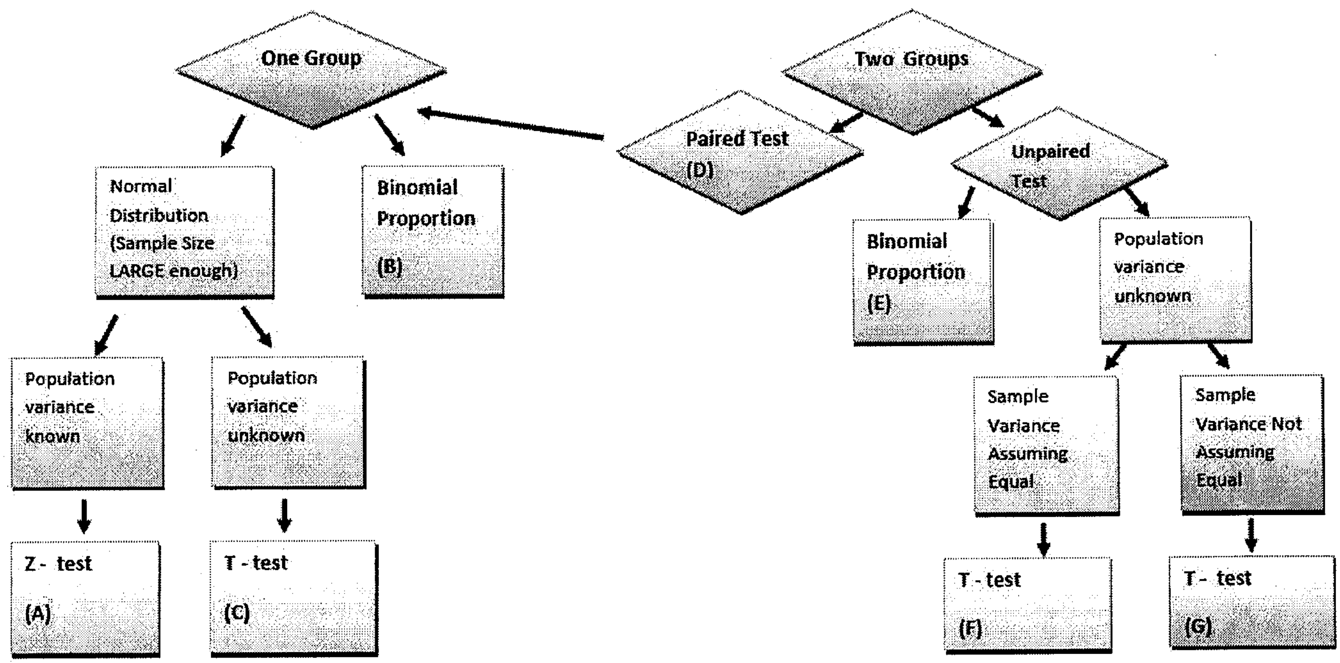
Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26039

You are previewing content.

http://online.wsj.com/article/SB119387051246678244.html
Have men and women's incomes gotten closer together since 1996? You want to determine this by asking a large sample of n_1 men and n_2 women for their annual income in 1996 and today. By comparing the difference in the mean income changes, we can tell whether the gap has gotten smaller by 1% or not. Assume that both male and female differences in incomes come from normal distributions such that $X_{1,i} \sim N(\mu_1, \sigma)$ and $X_{2,i} \sim N(\mu_2, \sigma)$.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question
Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem
The distribution of the population is ...
Comment on this question

Appendix A

Select one:

- ◉ Known
- ○ Unknown

Submit Answer

Correct!
The size of the population is ...
Comment on this question

Select one:

- ◉ Large
- ○ Small

Submit Answer

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- ◉ Measurement
- ○ Proportion

Submit Answer
No, try again

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26040

You are previewing content.

State the hypotheses that will help us determine whether there is a difference of between the two sample means is greater than 1%:

(a) H_0: $\overline{X}_1 - \overline{X}_2 = 0.01$

H_A: $\overline{X}_1 - \overline{X}_2 \neq 0.01$

(b) H_0: $\mu_1 - \mu_2 = 0.01$

H_A: $\mu_1 - \mu_2 \neq 0.01$

(c) H_0: $\mu_1 - \mu_2 \geq 0.01$

H_A: $\mu_1 - \mu_2 < 0.01$

(d) H_0: $\mu_1 - \mu_2 \leq 0.01$

H_A: $\mu_1 - \mu_2 > 0.01$

Comment on this question

Request Help

Select one:

- ○A
- ○B
- ○C
- ○D

Submit Answer

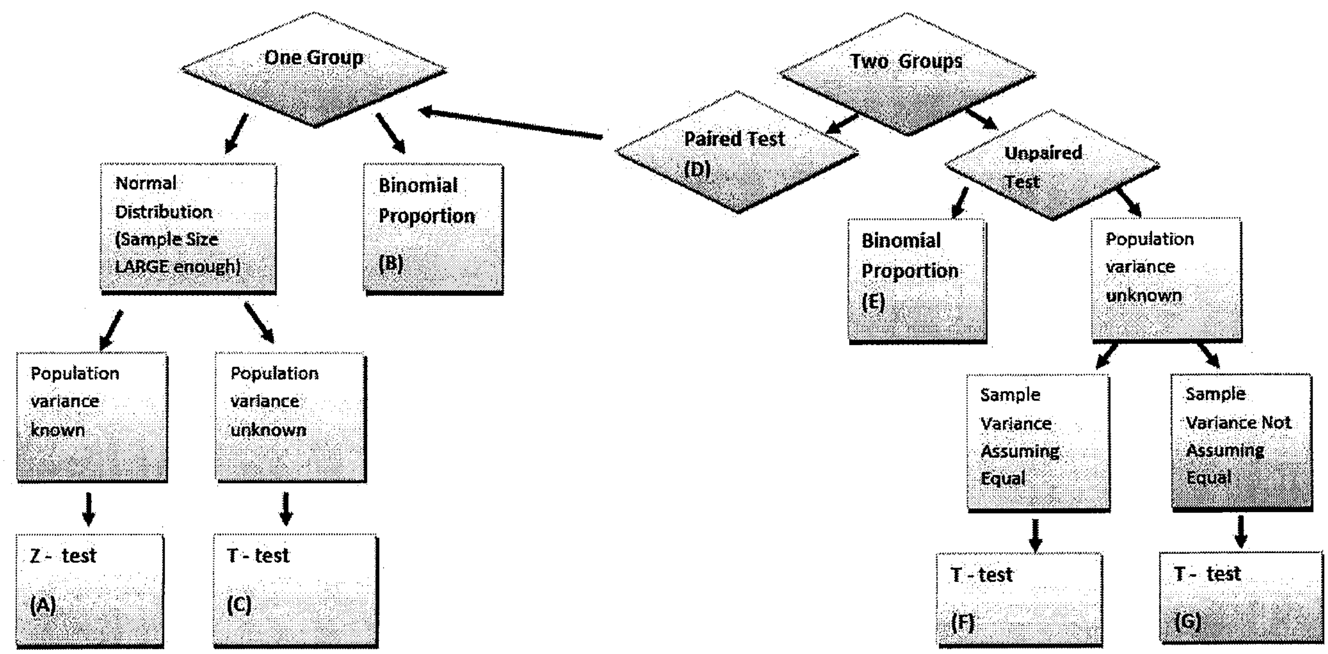
Appendix A

Assistment

Assistment #26019

You are previewing content.

A new Facebook feature called Beacon lets users broadcast their actions to others (*Thoughts on Beacon.* Mark Zuckerberg. http://blog.facebook.com/blog.php?post=7584397130). Some Facebook users are upset because the feature was too aggressive at sharing their information (*Facebook's Overblown Privacy Problems.* Andy Greenberg. http://www.forbes.com/technology/2007/12/05/facebook-beacon-opt-tech-internet-cx_ag_1205techfacebook.html). Suppose we would like to sample 714 Facebook users and determine whether at least 1.5% of them are worried about privacy.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question

Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem

The number of samples is ...

Appendix A

Comment on this question

Select one:

- ◉ One
- ○ Two

Submit Answer

Correct!
The size of the population is ...
Comment on this question

Select one:

- ◉ Large
- ○ Small

Submit Answer

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- ○ Measurement
- ◉ Proportion

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26021

You are previewing content.
Recall the problem above: we want to determine if more than 1.5% of Facebook users care about their privacy. Suppose a sample of 714 users finds that 1 cares about his privacy (joined a group advocating for it).

Perform a hypothesis test to see whether the proportion of users who care about their privacy is more than 1.5%, and report the p-value for this test below. Please provide your answer to four decimal places.
Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
Which test statistic should we use for this problem: z or t?
Comment on this question

Use the same guidelines for choosing a test statistic as you would for a confidence interval problem.
Comment on this hint

We're looking at the proportion of one sample.
Comment on this hint

For a hypothesis test of the proportion, if the population standard deviation is unknown, but the population is large, use the z statistic.
Comment on this hint

Select one:

- (•) z
- ( ) t

Submit Answer

Correct!
Calculate the test or sample value for t (to two decimal places).
Comment on this question

The expression for z* in this case is

$z^* = (\bar{x} - \alpha)/\sigma_{\bar{x}}$, where $\bar{x}$ is the sample proportion, α is the value we're testing for and $\sigma_{\bar{x}} = \sqrt{(p^*q/n)}$ is the standard error (also called the standard deviation of the sampling distribution of the mean.

Comment on this hint

The expression for t in this case is

$z^* = (\bar{x} - \alpha)/\sigma_{\bar{x}}$, where $\bar{x}$ is the sample proportion (.002), α is the value we're testing for (.001) and $\sigma_{\bar{x}} = \sqrt{(p^*q/n)}$ is the standard error ($\sqrt{(.015^*.985/714)} = 0.00454899$)

Comment on this hint

Request Help

Type your answer below:

-

Submit Answer

No, sorry

Assistment

Assistment #26020

You are previewing content.

A new Facebook feature called Beacon lets users broadcast their actions to others (*Thoughts on Beacon.* Mark Zuckerberg. http://blog.facebook.com/blog.php?post=7584397130). Some Facebook users are upset because the feature was too aggressive at sharing their information (*Facebook's Overblown Privacy Problems*. Andy Greenberg. http://www.forbes.com/technology/2007/12/05/facebook-beacon-opt-tech-internet-cx_ag_1205techfacebook.html). Suppose we would like to sample 714 Facebook users and determine whether at least 1.5% of them are worried about privacy.

State the hypotheses that will help us determine whether at least 1.5% of Facebook users are worried about their privacy.

(a) H_0: $\overline{X} = 0.015$

H_A: $\overline{X} \neq 0.015$

(b) H_0: $\mu = 0.015$

H_A: $\mu \neq 0.015$

(c) H_0: $\mu \geq 0.015$

H_A: $\mu < 0.015$

(d) H_0: $\mu \leq 0.015$

H_A: $\mu > 0.015$

Comment on this question

There are two hypotheses in a hypothesis test: the null hypothesis and the alternative hypothesis. Hypothesis tests are always about population parameters.

Comment on this hint

We believe the null hypothesis unless there is significant enough evidence to convince us to reject it in favor of the alternative hypothesis.

Comment on this hint

H_0: $\mu \leq 0.001$

H_A: $\mu > 0.001$

Comment on this hint

Select one:

- ○ A
- ○ B
- ○ C
- ◉ D

Submit Answer

Sorry, that's not correct. The problem asks you to test whether the population proportion is above a value. Put another way, we're interested in knowing if the population proportion is now greater than some number. The alternative hypothesis here test's to see if the population proportion is smaller than a number.

Correct!

You are done with this problem!

Comment on this problem

Assistment

Assistment #26022

You are previewing content.
Recall the problem above: we want to determine if more than 1.5% of Facebook users care about their privacy. Suppose a sample of 714 users finds that 1 cares about his privacy (joined a group advocating for it).

Consider the hypothesis test to see whether the proportion of Facebook users who care about their privacy is greater than 1.5%. You determined the appropriate hypotheses to be:

H_0: $\mu \leq 0.015$

H_A: $\mu > 0.015$

You determined the p-value for this two sided test to be .9986; Which of the following statements is true?

Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint
If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.
Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ◯ The proportion of Facebook users who care about their privacy is less than 1.5%
- ◯ The proportion of Facebook users who care about their privacy is more than 1.5%
- ◉ There is not enough information to draw either of the above conclusions.

Submit Answer
No, try again

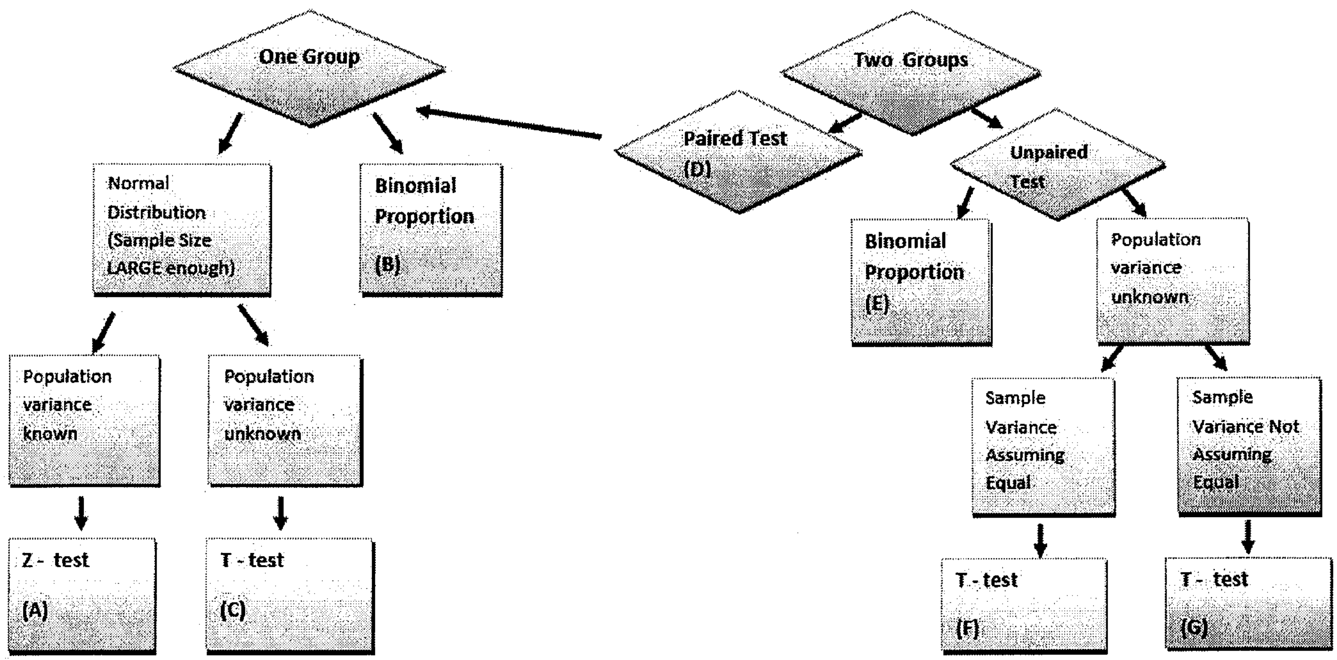
Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26039

You are previewing content.

http://online.wsj.com/article/SB119387051246678244.html
Have men and women's incomes gotten closer together since 1996? You want to determine this by asking a large sample of n_1 men and n_2 women for their annual income in 1996 and today. By comparing the difference in the mean income changes, we can tell whether the gap has gotten smaller by 1% or not. Assume that both male and female differences in incomes come from normal distributions such that $X_{1,i} \sim N(\mu_1, \sigma)$ and $X_{2,i} \sim N(\mu_2, \sigma)$.



Using the chart above, what type of test should be used to calculate this p-value?

Comment on this question
Request Help

Select one:

- A
- B
- C
- D
- E
- F
- G

Submit Answer

Let's move on and figure out this problem
The distribution of the population is ...
Comment on this question

Appendix A

Select one:

- Known
- Unknown

Submit Answer

Correct!
The size of the population is ...
Comment on this question

Select one:

- Large
- Small

Submit Answer
No, sorry

Correct!
Are we looking at a measurement or a proportion?
Comment on this question

Select one:

- Measurement
- Proportion

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26040

You are previewing content.

State the hypotheses that will help us determine whether there is a difference of between the two sample means is greater than 1%:

(a) H_0: $\overline{X}_1 - \overline{X}_2 = 0.01$

H_A: $\overline{X}_1 - \overline{X}_2 \neq 0.01$

(b) H_0: $\mu_1 - \mu_2 = 0.01$

H_A: $\mu_1 - \mu_2 \neq 0.01$

(c) H_0: $\mu_1 - \mu_2 \geq 0.01$

H_A: $\mu_1 - \mu_2 < 0.01$

(d) H_0: $\mu_1 - \mu_2 \leq 0.01$

H_A: $\mu_1 - \mu_2 > 0.01$

Comment on this question

There are two hypotheses in a hypothesis test: the null hypothesis and the alternative hypothesis. Hypothesis tests are always about population parameters.

Comment on this hint

We believe the null hypothesis unless there is significant enough evidence to convince us to reject it in favor of the alternative hypothesis.

Comment on this hint

H_0: $\mu_1 - \mu_2 \leq 0.01$

H_A: $\mu_1 - \mu_2 > 0.01$

Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ◉D

Submit Answer

Sorry, that's not correct. The problem asks you to test whether the difference between two means is above a value. Put another way, we're interested in knowing if the difference between two means is larger than some number. The alternative hypothesis here test's to see if the difference between two means is smaller than a number.

Correct!

You are done with this problem!

Comment on this problem

Assistment

Assistment #26042

You are previewing content.
Recall the problem above: we want to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.

The mean percentage change in male income was 2.7% while the mean percentage change in female income was 7.1%. The sample standard deviation of this difference is 2.2%.

Consider the hypothesis test to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%. You determined the appropriate hypotheses to be:

$H_0: \mu_1 - \mu_2 \leq 0.01$

$H_A: \mu_1 - \mu_2 > 0.01$

You determined the p-value for this two sided test to be .1824. Which of the following statements is true?

Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.
Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.
- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is not greater than 1%.
- ◉ There is not enough information to draw either of the above conclusions.

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26042

You are previewing content.
Recall the problem above: we want to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.

The mean percentage change in male income was 2.7% while the mean percentage change in female income was 7.1%. The sample standard deviation of this difference is 2.2%.

Consider the hypothesis test to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%. You determined the appropriate hypotheses to be:

$H_0: \mu_1 - \mu_2 \leq 0.01$

$H_A: \mu_1 - \mu_2 > 0.01$

You determined the p-value for this two sided test to be .1824. Which of the following statements is true?

Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.
Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.
- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is not greater than 1%.
- ◉ There is not enough information to draw either of the above conclusions.

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26042

You are previewing content.
Recall the problem above: we want to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.

The mean percentage change in male income was 2.7% while the mean percentage change in female income was 7.1%. The sample standard deviation of this difference is 2.2%.

Consider the hypothesis test to determine if the difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%. You determined the appropriate hypotheses to be:

$H_0: \mu_1 - \mu_2 \leq 0.01$

$H_A: \mu_1 - \mu_2 > 0.01$

You determined the p-value for this two sided test to be .1824. Which of the following statements is true?

Comment on this question

A small p-value indicates that the null hypothesis should be rejected.
Comment on this hint

If the p-value is less than the level of significance, we reject the null hypothesis in favor of the alternative.
Comment on this hint
Here the p-value is greater than the level of significance, so H_0 should not be rejected in favor of H_A at the .05 level.
Comment on this hint

Select one:

- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is greater than 1%.
- ○ The difference between the mean percentage change in male income and the mean percentage change in female income is not greater than 1%.
- ◉ There is not enough information to draw either of the above conclusions.

Submit Answer

Correct!
You are done with this problem!
Comment on this problem

Assistment

Assistment #26637

You are previewing content.
Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129.

Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a) What is the probability that both individuals are uninsured?

A. 0.871
B. 0.016641
C. 0.258
D. 0.241359
E. 0.024

Comment on this question
The probability for the selected man and woman to be uninsured is the same. P(woman uninsured)=P (man uninsured)=0.129.
Comment on this hint
For any randomly selected people, the probability for her or him to be uninsured is independent from any other person. For independent events A and B, $P(A \cap B) = P(A)*P(B)$
Comment on this hint
By the property of indepence, P(both uninsured)=P(woman uninsured) * P(man uninsured)= 0.129*0.129= 0.016641
Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Submit Answer

Assistment

Assistment #26648

You are previewing content.

Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129. Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(b)What is the probability that both adults have health insurance coverage?

A. 0.016641
B. 0.241359
C. 0.983359
D. 0.758641
E. 0.112359

Comment on this question

The probability for the selected woman and man to be insured is same. P(woman insured)=P(man insured)=1-0.129=0.871

Comment on this hint

For any randomly selected people, the probability for her or him to be insured is independent from any other person. For independent events A and B, $P(A \cap B) = P(A)*P(B)$

Comment on this hint

By the property of indepence, P(both insured)=P(woman insured) * P(man insured)= (1-0.129)*(1-0.129)=0.758641

Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Submit Answer

Assistment

Assistment #26659

You are previewing content.
Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129.

c) If five unrelated adults between the ages of 45 and 64 are chosen from the population, what is the probability that all five are uninsured?

A. 3.5723051649e-005
B. 0.501292001353351
C. 0.645
D. 0.000276922881
E. 4.608273662721e-006

Comment on this question
The probability for the five selected persons to be uninsured is same. P(1st uninsured)=P(2nd uninsured) =P(3rd uninsured)=P(4th uninsured)=P(5th uninsured)=0.129
Comment on this hint
As five adults are unrelated, the probability for any one of them to be uninsured is independent from any other persons. For independent events A and B, P(A∩B) = P(A)*P(B)
Comment on this hint
By the property of indepence, P(five uninsured)=P(1st uninsured)*P(2nd uninsured)*P(3rd uninsured) *P(4th uninsured)*P(5th uninsured)= 0.129*0.129*0.129*0.129*0.129=3.5723051649e-005
Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Submit Answer

Assistment

Assistment #26665

You are previewing content.
Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.124.

Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a)What is the probability that both individuals are uninsured?

A. 0.876
B. 0.015376
C. 0.248
D. 0.232624
E. 0.02

Comment on this question
Request Help

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Submit Answer

Let's move on and figure out this problem
Let's look at the solution for a problem similar to the one above:

Suppose that recent survey data conducted by the music industry indicates that the probability that a randomly chosen college student has purchased a CD in the last 12 months is 0.26.

Now suppose that you randomly select a student from one college and an unrelated student from another college. What is the probability that both have purchased a CD in the last 12 months?

Solution:

Let A be the event that the first college student purchased a CD in the last 12 months.
Let B be the event that the second college student purchased a CD in the last 12 months.

In mathematical terms, the problem asks for P(A∩B).

As the two students are from different colleges and unrelated, we can presume that the two events are independent. We know that for independent events A and B, P(A∩B) = P(A)*P(B).

Here P(A) = 0.26 and likewise P(B) = 0.26.

P(A∩B) = P(A)*P(B) = 0.26 * 0.26 = .0676.

This problem is equivalent to the case of flipping a fair coin twice and asking what the probability is of getting two heads.

Let A be the event of getting a head on the first toss.
Let B be the event of getting a head on the second toss.

P(A∩B) = P(A)*P(B) = 0.5 * 0.5 = .025.

Comment on this question

Select one:

- I have read through the worked example and now I am ready to answer the original question

Submit Answer

Correct!
Now try the original problem again.

You may look back at the worked example if that helps you.

Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.124.
Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a)What is the probability that both individuals are uninsured?

A. 0.876
B. 0.015376
C. 0.248
D. 0.232624
E. 0.02

Comment on this question

Select one:

Appendix A

- ◯A
- ◯B
- ◯C
- ◯D
- ◯E

Submit Answer

Assistment

Assistment #26933

You are previewing content.

In the Netherlands, healthy females between the ages of 25 and 39 have a distribution of serum uric acid levels that is approximately normal with mean $\mu = 348$ µmol/l and standard deviation $\sigma = 78$ µmol/l.

A) Rounding to three digits after the decimal point, What proportion of the females have a serum uric acid level between 300 and 400 µmol/l?

Comment on this question

Type your answer below (mathematical expression):

-

Submit Answer

No, sorry

Assistment

Assistment #26943

You are previewing content.

In the Netherlands, healthy females between the ages of 25 and 39 have a distribution of serum uric acid levels that is approximately normal with mean $\mu = 348$ µmol/l and standard deviation $\sigma = 78$ µmol/l.

B) Rounding to three digits after the decimal point, What proportion of samples of size 5 have a mean serum uric acid level between 300 and 400 µmol/l?

Comment on this question

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26973

You are previewing content.

In the Netherlands, healthy females between the ages of 25 and 39 have a distribution of serum uric acid levels that is approximately normal with mean $\mu = 348$ µmol/l and standard deviation $\sigma = 78$ µmol/l.

C) Rounding to three digits after the decimal point, What proportion of samples of size 10 have a mean serum uric acid level between 300 and 400 µmol/l?

Comment on this question

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26953

You are previewing content.

In the Netherlands, healthy females between the ages of 25 and 39 have a distribution of serum uric acid levels that is approximately normal with mean $\mu = 348$ μmol/l and standard deviation $\sigma = 78$ μmol/l.

D) If you were to construct an interval that encloses 95% of the means of samples of size 10, which would be shorter, a symmetric interval or an asymmetric one?

Comment on this question

Select one:

- ○ A symmetric interval
- ○ An asymmetric interval

Submit Answer

Assistment

Assistment #26872

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

A) Rounding to three digits after the decimal point, What is the probability that no suicides will be reported during a given month?

Comment on this question

The probability of X suicides being reported during a given month (assuming that suicides follow a Poisson distribution with mean 2.25 suicides per month) is (e^-2.25)*(2.25^X)/X!

Comment on this hint

We are looking for the probability that no suicides are reported. Thus, find the probability that 0 suicides are reported during a given month using the formula above.

Comment on this hint

Evaluting (e^-2.25)*(2.25^X)/X! for X = 0,
(e^-2.25)*(2.25^X)/X!
= (e^-2.25)*(2.25^0)/0!
= e^-2.25
= 0.105399224561864 ≈ 0.105

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26882

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

B) Rounding to three digits after the decimal point, What is the probability that at most four suicides will be reported?

Comment on this question

The probability of X suicides being reported during a given month (assuming that suicides follow a Poisson distribution with mean 2.25 suicides per month) is (e^-%v{a) *(%v{a^X) / X!

Comment on this hint

We are looking for the probability that no more than four suicides are reported. Our formula applies in finding the exact probability of X suicides being reported. Thus, to find the probability that no more than four suicides are reported, find the probability that X = 0, 1, 2, 3, and 4 suicides are reported during a given month using the formula above and take the sum of the probabilities of these independent events.

Comment on this hint

P(at most four) = P(none) + P(one) + P(two) + P(three) + P(four)
= (e^-%v{a) *(%v{a^0) / 0! + (e^-%v{a) *(%v{a^1) / 1! +(e^-%v{a) *(%v{a^2) / 2! + (e^-%v{a) *(%v {a^3) / 3! + (e^-%v{a) *(%v{a^4) / 4!
= 0.921985892590723 ≈ 0.922

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26892

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

C) Rounding to three digits after the decimal point, What is the probability that six or more suicides will be reported?

Comment on this question

The probability of X suicides being reported during a given month (assuming that suicides follow a Poisson distribution with mean 2.25 suicides per month) is (e^-2.25)*(2.25^X)/X!

Comment on this hint

We are looking for the probability that six or more suicides are reported. We only have a formula for finding the exact probability of X suicides being reported. Thus, to find the probability that six or more suicides are reported, find the probability that less than six suicides are reported and subtract this from 1.

Comment on this hint

P(six or more) = P(less than six)
= 1 - (P(none) + P(one) + P(two) + P(three) + P(four) + P(five))
= 1-((e^-%v{a) *(%v{a^0) / 0! + (e^-%v{a) *(%v{a^1) / 1! +(e^-%v{a) *(%v{a^2) / 2! + (e^-%v{a) * (%v{a^3) / 3! + (e^-%v{a) *(%v{a^4) / 4! + (e^-%v{a) *(%v{a^5) / 4!)
= 1 - 0.972634645936698 ≈ 0.027

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26902

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

A) Rounding to three digits after the decimal point, What is the probability that no suicides will be reported during a given month?

Comment on this question

Worked Example

Suppose the number of silly t-shirts being reported at MIT during a given day follows a Poisson distribution with mean 17. What is the probability that 16 silly t-shirts are reported during a given day?

P(X=16) = (e^-17)*(17^16)/16! = 0.0962846277984453 ≈ 0.096

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26912

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

B) Rounding to three digits after the decimal point, What is the probability that at most four suicides will be reported?

Comment on this question

Worked Example:

Suppose the number of silly t-shirts being reported at MIT during a given day follows a Poisson distribution with mean 17. What is the probability that between 5 and 7 silly t-shirts are reported during a given day?

P(between 5 and 7 silly t-shirts)
= P(5 silly t-shirts) + P(6 silly t-shirts) + P(7 silly t-shirts)
=(e^-17)*(17^5)/5! + (e^-17)*(17^6)/6! + (e^-17)*(17^7)/7!
=0.0052483210264l174 ≈ 0.005

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26922

You are previewing content.

In a particular county, the average number of suicides reported each month is 2.25. Assume that the number of suicides follows a Poisson distribution.

C) Rounding to three digits after the decimal point, What is the probability that six or more suicides will be reported?

Comment on this question

Worked Example:

Suppose the number of silly t-shirts being reported at MIT during a given day follows a Poisson distribution with mean 17. What is the probability that more than three silly t-shirts are reported during a given day?

P(more than 3 silly t-shirts) = 1 - P(less than 3 silly t-shirts)
=1 - (P(0 silly t-shirts) + P(1 silly t-shirt) + P(2 silly t-shirts))
=1 - ((e^-17)*(17^0)/0! + (e^-17)*(17^1)/1! + (e^-17)*(17^2)/2!)
=1 - 4.05851894364906e-05
=0.999959414810563
≈ 1

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #27007

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X P(X = x)
0 0.647
1 0.161
2 0.061
3 0.019
4 0.014
5 0.098
Total 1.0

What is the probability that a child receives exactly three diagnostic services during an office visit to a pediatric specialist?

Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
Let X be a discrete random variable that represents the number of vaccines a child receives from birth to 12 years old. The probability distribution for X appears below.
X P(x)
0 0.213
1 0.137
2 0.457
3 0.006
4 0.009
5 0.178
Total 1.0
The probability that a child receives exactly 5 vaccines is 0.178.

Comment on this question

Select one:

- ◉ I have read the worked example.

Submit Answer

Correct!
Comment on this question

Type your answer below:

-

Submit Answer

Assistment

Assistment #27029

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X	P(X = x)
0	0.647
1	0.161
2	0.061
3	0.019
4	0.014
5	0.098
Total	1.0

What is the probability that he or she receives at least one service?

Comment on this question

Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
Let X be a discrete random variable that represents the number of vaccines a child receives from birth to 12 years old. The probability distribution for X appears below.

X	P(x)
0	0.213
1	0.137
2	0.457
3	0.006
4	0.009
5	0.178
Total	1.0

The probability that a child receives at least 3 vaccines is 0.187.

Comment on this question

Select one:

- I have read the worked example.

Submit Answer

Correct!
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X P(X = x)
0 0.647
1 0.161
2 0.061
3 0.019
4 0.014
5 0.098
Total 1.0

What is the probability that he or she receives at least one service?

Comment on this question

Type your answer below:

-

Submit Answer

Assistment

Assistment #27051

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X P(X = x)
0 0.647
1 0.161
2 0.061
3 0.019
4 0.014
5 0.098
Total 1.0

What is the probability that the child receives exactly three services given that he or she receives at least one service? (Round to three digits)

Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
Let X be a discrete random variable that represents the number of vaccines a child receives from birth to 12 years old. The probability distribution for X appears below.
X P(x)
0 0.213
1 0.137
2 0.457
3 0.006
4 0.009
5 0.178
Total 1.0
What is the probability that a child receives at least 4 vaccines given that he or she receives at least 2 vaccines?

$P(x \geq 4 | x \geq 2) = (P(\{x \geq 4\} \cap \{x \geq 2\}))/(P(x \geq 2)) = (P(x \geq 4))/(P(x \geq 2))$

Comment on this question

Select one:

- I have read the worked example.

Submit Answer

Correct!
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X	P(X = x)
0	0.647
1	0.161
2	0.061
3	0.019
4	0.014
5	0.098
Total	1.0

What is the probability that the child receives exactly three services given that he or she receives at least one service? (Round to three digits)

Comment on this question
Request Help

Type your answer below:

-

Submit Answer

Assistment

Assistment #26996

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X	P(X = x)
0	0.647
1	0.161
2	0.061
3	0.019
4	0.014
5	0.098
Total	1.0

What is the probability that a child receives exactly three diagnostic services during an office visit to a pediatric specialist?

Comment on this question

P(X=x) denotes the probability that a child receives exactly x diagnostic services.

Comment on this hint

Type your answer below:

-

Submit Answer

Assistment

Assistment #27018

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X P(X = x)
0 0.647
1 0.161
2 0.061
3 0.019
4 0.014
5 0.098
Total 1.0

What is the probability that he or she receives at least one service?

Comment on this question

P(X≥x) denotes the probability that a child receives at least x diagnostic services.

Comment on this hint

In discrete cases, P(X≥1) = ∑P(k) = P(1) + P(2) + P(3) + P(4) + P(4+)

Comment on this hint

Or we may also simplify the problem by thinking the other way: P(X≥1) = 1 – P(X<1) = 1 – P(0)

Comment on this hint

Type your answer below:

-

Submit Answer

Assistment

Assistment #27040

You are previewing content.
Let X be a discrete random variable that represents the number of diagnostic services a child receives during an office visit to a pediatric specialist; these services include procedures such as blood tests and urinanalysis. The probability distribution for X appears below.

X	P(X = x)
0	0.647
1	0.161
2	0.061
3	0.019
4	0.014
5	0.098
Total	1.0

What is the probability that the child receives exactly three services given that he or she receives at least one service? (Round to three digits)

Comment on this question

$P(A|x \geq B)$ denotes the probability that a child receives exactly A diagnostic services given that he or she receives at least B services.

Comment on this hint

$P(p|q) = (P(p \cap q))/(P(q))$

Comment on this hint

$P(3|x \geq 1) = (P(\{x=3\} \cap \{x \geq 1\}))/(P(x \geq 1)) = (P(x=3))/(P(x \geq 1))$

Comment on this hint

Type your answer below:

-

Submit Answer

Assistment

Assistment #27077

You are previewing content.
Consider a group of 9 individuals selected from the population of 65- to 74- year-olds in the United States. The number of persons in this sample who suffer from diabetes is a binomial random variable with parameters n=9 and p =0.211.

Without regard to order, in how many ways can you select four individuals from this group of n?

Comment on this question

Without order and without replacement, select four from n is simply "n choose four".

Comment on this hint

Type your answer below:

-

Submit Answer

Assistment

Assistment #27099

You are previewing content.
Consider a group of 9 individuals selected from the population of 65- to 74- year-olds in the United States. The number of persons in this sample who suffer from diabetes is a binomial random variable with parameters n=9 and p =0.211.

What is the probability that exactly two of the individuals in the sample suffer from diabetes?(Round your answer to two decimal places)

Comment on this question

Request Help

Type your answer below:

-

Submit Answer

Let's move on and figure out this problem
According to the National Health Survey, 9.8% of the population of 18- to 24- year-olds in the United States are left-handed.

What is the probability that exactly three out of ten persons are left-handed?

The probability that exactly 3 out of 10 individuals chosen with probability = 9.8% for each individual is

$P(3) = \binom{10}{3} 9.8\%^3 (1 - 9.8\%)^7$

Comment on this question

Select one:

- I have read the worked example.

Submit Answer

Correct!

Consider a group of 9 individuals selected from the population of 65- to 74- year-olds in the United States. The number of persons in this sample who suffer from diabetes is a binomial random variable with parameters n=9 and p =0.211.

What is the probability that exactly two of the individuals in the sample suffer from diabetes?(Round your answer to two decimal places)

Comment on this question

Type your answer below:

-

Submit Answer

Assistment

Assistment #27055

You are previewing content.
Consider a group of 9 individuals selected from the population of 65- to 74- year-olds in the United States. The number of persons in this sample who suffer from diabetes is a binomial random variable with parameters n=9 and p =0.211.

If you wish to make a list of the 9 persons chosen, in how many ways can they be ordered?

Comment on this question

For the first place, we could choose any one of the n person. So there are n ways for the first place. For the second place, we could choose any one from the rest 6 persons. So for the second place, we have n-1 choices, and so on.

Comment on this hint

Since we have n places and n persons, we have n! ways to order n persons.

Comment on this hint

Type your answer below:

-

Submit Answer

Assistment

Assistment #26743

You are previewing content.

Among females in the United States between 18 and 74 years of age, diastolic blood pressure is normally distributed with mean μ=81 mm Hg and standard deviation σ=12.5 mm Hg.
(a) What is the probability that a randomly selected woman has a diastolic blood pressure less than 70 mm Hg? (round to 2 decimal places)

Comment on this question

The female diastolic blood pressure is normally distributed.

Comment on this hint

The mean for this normal distribution is μx=81 and the standard deviation is σx=12.5.

Comment on this hint

The probability that a randomly selected woman has a diastolic blood pressure less than 70 mm Hg is P (X<70). Applying Z-transformation, $P(X < 70) = P[(X-\mu x)/\sigma x < (70-\mu x)/\sigma x]=P[Z < (70-81)/12.5]= 0.19$

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26843

You are previewing content.
In Denver, Colorado, the distribution of daily measures of ambient nitric acid-a corrosive liquid-is skewed to the right; it has mean μ= 1.81 μg/m3 and standard deviation σ=2.4 μg/m3. What is the distribution of means of samples of size 37 selected from the population? Calculate the mean and standard deviation for the sample mean.

Comment on this question

The mean for the sample mean of size 37 is μ= 1.81, which is the same with the population mean. The standard deviation for sample mean of size 37 is $\sigma s = \sigma/\sqrt{n} = 2.4/\sqrt{37} = 0.395$.

Comment on this hint

Select one:

- ○ A. normal distribution / mean=1.81 / standard deviation=0.445
- ○ B. normal distribution / mean=1.81 / standard deviation=0.345
- ○ C. normal distribution / mean=1.82 / standard deviation=0.395
- ○ D. normal distribution / mean=1.8 / standard deviation=0.395
- ○ E. normal distribution / mean=1.81 / standard deviation=0.395
- ○ F. binomial distribution / mean=1.81 / standard deviation=0.345
- ○ G. binomial distribution / mean=1.81 / standard deviation=0.445
- ○ H. binomial distribution / mean=1.82 / standard deviation=0.395
- ○ I. binomial distribution / mean=1.8 / standard deviation=0.395
- ○ J. binomial distribution / mean=1.81 / standard deviation=0.395

Submit Answer

Assistment

Assistment #26637

You are previewing content.
Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129.

Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a) What is the probability that both individuals are uninsured?

A. 0.871
B. 0.016641
C. 0.258
D. 0.241359
E. 0.024

Comment on this question
The probability for the selected man and woman to be uninsured is the same. P(woman uninsured)=P (man uninsured)=0.129.
Comment on this hint
For any randomly selected people, the probability for her or him to be uninsured is independent from any other person. For independent events A and B, $P(A \cap B) = P(A)*P(B)$
Comment on this hint
By the property of indepence, P(both uninsured)=P(woman uninsured) * P(man uninsured)= 0.129*0.129= 0.016641
Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D
- ○E

Submit Answer

Assistment

Assistment #26670

You are previewing content.
Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129.

Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a)What is the probability that both individuals are uninsured?

A. 0.871
B. 0.016641
C. 0.258
D. 0.241359
E. 0.02

Comment on this question
Request Help

Select one:

- A
- B
- C
- D
- E

Submit Answer

Let's move on and figure out this problem
Let's look at the solution for a problem similar to the one above:

Suppose that recent survey data conducted by the music industry indicates that the probability that a randomly chosen college student has purchased a CD in the last 12 months is 0.26.

Now suppose that you randomly select a student from one college and an unrelated student from another college. What is the probability that both have purchased a CD in the last 12 months?

Solution:

Let A be the event that the first college student purchased a CD in the last 12 months.
Let B be the event that the second college student purchased a CD in the last 12 months.

In mathematical terms, the problem asks for P(A∩B).

As the two students are from different colleges and unrelated, we can presume that the two events are independent. We know that for independent events A and B, P(A∩B) = P(A)*P(B).

Here P(A) = 0.26 and likewise P(B) = 0.26.

P(A∩B) = P(A)*P(B) = 0.26 * 0.26 = .0676.

This problem is equivalent to the case of flipping a fair coin twice and asking what the probability is of getting two heads.

Let A be the event of getting a head on the first toss.
Let B be the event of getting a head on the second toss.

P(A∩B) = P(A)*P(B) = 0.5 * 0.5 = .025.

Comment on this question

Select one:

- I have read through the worked example and now I am ready to answer the original question

Submit Answer

Correct!
Now try the original problem again.

You may look back at the worked example if that helps you.

Looking at the United States population in 1993, the probability that an adult between the ages of 45 and 64 does not have health insurance coverage of any kind is 0.129.
Suppose that you randomly select a 47-year-old woman and an unrelated 59-year-old man from this population.

(a)What is the probability that both individuals are uninsured?

A. 0.871
B. 0.016641
C. 0.258
D. 0.241359
E. 0.02

Comment on this question

Select one:

Appendix A

- ○ A
- ○ B
- ○ C
- ○ D
- ○ E

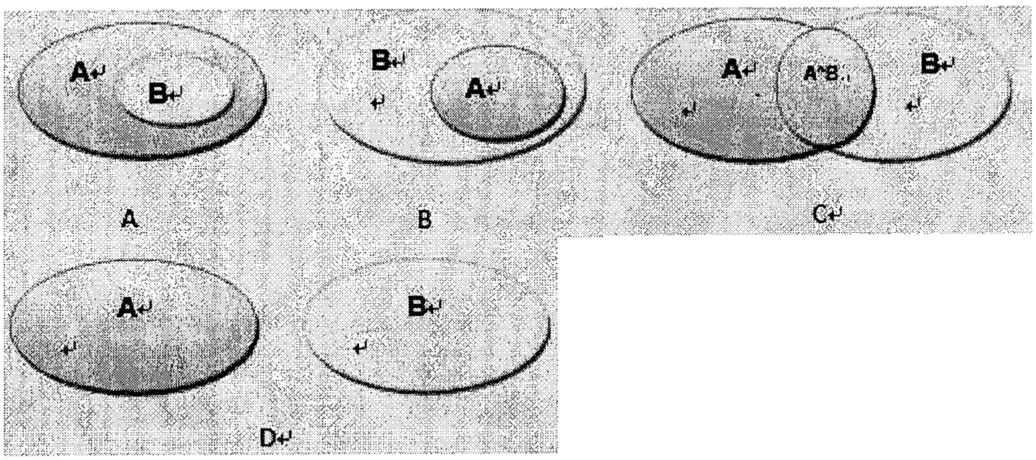
Submit Answer

Assistment

Assistment #26558

You are previewing content.
For Mexican American infants born in Arizona in 1986 and 1987, the probability that a child's gestational age is less than 37 weeks is 0.133 and the probability that his or her birth weight is less than 2500 grams is 0.052. Furthermore, the probability that these two events occur simultaneously is 0.023.

Let A be the event that an infant's gestational age is less than 37 weeks and B the event that his or her birth weight is less than 2500 grams. Select a Venn diagram from the following to illustrate the relationship between events A and B.



Comment on this question

A recent survey of students at a major university indicated that the probability of a student seeing on average more than five movies per week is 0.168. The same survey showed that the probability of a student having a part time job is 0.22 and that the probability a student sees more than five movies per week and has a part time job is 0.043. Let A be the event that a student at the university sees more than five movies per week and B the event that a student at the university has a part time job. Which of the following Venn diagrams best represents the relationship between A and B? You can eliminate answer (d) as for this diagram P $(A \cap B) = 0$ (as circles A and B do not intersect), while the problem specifies that P $(A \cap B) = 0.043$. You can eliminate answer (a) as this diagram indicates all students who have part time jobs see more than five movies per week. In mathematical terms, this would mean $P(A|B) = 1$. The problem gives no indication that this is the case, and in fact we can determine from the data that P $(A|B) = = = 0.195$. Likewise, you can eliminate answer (b) as this diagram indicates all students who see more than five movies per week have part time jobs. In mathematical terms, this would mean $P(A|B) = 1$. The problem gives no indication that this is the case, and in fact we can determine from the data that P $(B|A) = = = 0.256$. Answer (c) is correct as it takes into account all possibilities described by the problem: A student who sees more than five movies per week and who has a part time job (the intersection of the two circles A and B). A student who does not see more than five movies per week and who has a part time job (circle B except for where it intersects with circle A). A student who sees more than five movies per week and who does not have a part time job (circle A except for where it intersects with circle B). A student who does not see more than five movies per week and who does not have a part time job (the rectangle not covered by the circles).

Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D

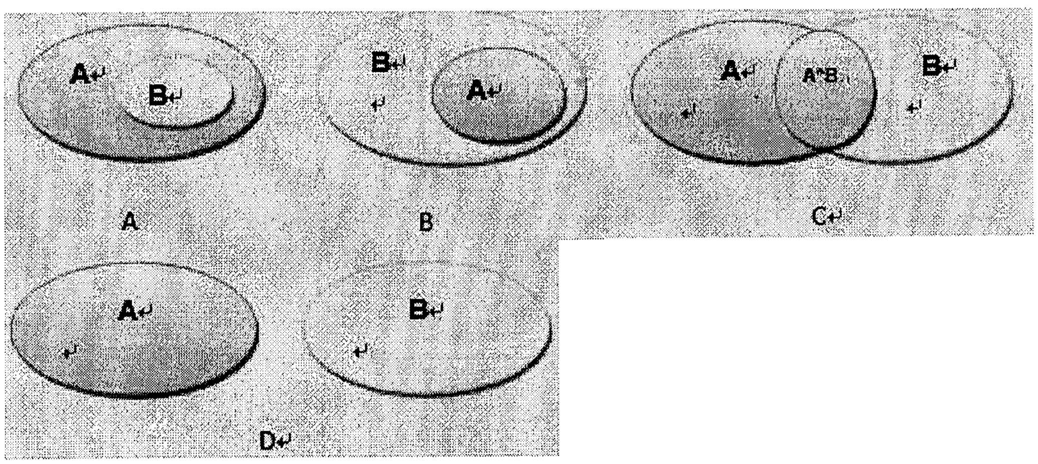
Submit Answer

Assistment

Assistment #26547

You are previewing content.
For Mexican American infants born in Arizona in 1986 and 1987, the probability that a child's gestational age is less than 37 weeks is 0.133 and the probability that his or her birth weight is less than 2500 grams is 0.052. Furthermore, the probability that these two events occur simultaneously is 0.023.

Let A be the event that an infant's gestational age is less than 37 weeks and B the event that his or her birth weight is less than 2500 grams. Select a Venn diagram from the following to illustrate the relationship between events A and B.



Comment on this question

The Venn diagram should include all possibilities.

Comment on this hint

There are some babies whose gestational age is less than 37 weeks (A), there are some babies whose birth weight is less than 2500 grams (B), and there are some babies whose gestational age is less than 37 weeks and whose birth weight is less than 2500 grams ($A \cap B$).

Comment on this hint

Select one:

- ○A
- ○B
- ○C
- ○D

Submit Answer

Assistment

Assistment #26854

You are previewing content.

In a particular county, the average number of suicides reported each month is %v{a}. Assume that the number of suicides follows a Poisson distribution.

A) Rounding to three digits after the decimal point, What is the probability that no suicides will be reported during a given month?

Comment on this question

The probability of X suicides being reported during a given month (assuming that suicides follow a Poisson distribution with mean %v{a} suicides per month) is (e^-%v{a})*(%v{a}^X)/X!

Comment on this hint

We are looking for the probability that no suicides are reported. Thus, find the probability that 0 suicides are reported during a given month using the formula above.

Comment on this hint

Evaluting (e^-%v{a})*(%v{a}^X)/X! for X = 0,
(e^-%v{a})*(%v{a}^X)/X!
= (e^-%v{a})*(%v{a}^0)/0!
= e^-%v{a}
= %v{Math.exp(-a)} ≈ %v{(Math.exp(-a)*a**0*1000.0).round()/1000.0}

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26855

You are previewing content.

In a particular county, the average number of suicides reported each month is %v{a}. Assume that the number of suicides follows a Poisson distribution.

B) Rounding to three digits after the decimal point, What is the probability that at most four suicides will be reported?

Comment on this question

The probability of X suicides being reported during a given month (assuming that suicides follow a Poisson distribution with mean %v{a} suicides per month) is (e^-%v{a) *(%v{a^X) / X!

Comment on this hint

We are looking for the probability that no more than four suicides are reported. Our formula applies in finding the exact probability of X suicides being reported. Thus, to find the probability that no more than four suicides are reported, find the probability that X = 0, 1, 2, 3, and 4 suicides are reported during a given month using the formula above and take the sum of the probabilities of these independent events.

Comment on this hint

P(at most four) = P(none) + P(one) + P(two) + P(three) + P(four)
= (e^-%v{a) *(%v{a^0) / 0! + (e^-%v{a) *(%v{a^1) / 1! +(e^-%v{a) *(%v{a^2) / 2! + (e^-%v{a) *(%v {a^3) / 3! + (e^-%v{a) *(%v{a^4) / 4!
= %v{sum} ≈ %v{sum_r}

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26858

You are previewing content.

In a particular county, the average number of suicides reported each month is %v{a}. Assume that the number of suicides follows a Poisson distribution.

B) Rounding to three digits after the decimal point, What is the probability that at most four suicides will be reported?

Comment on this question

Worked Example:

Suppose the number of silly t-shirts being reported at MIT during a given day follows a Poisson distribution with mean 17. What is the probability that between 5 and 7 silly t-shirts are reported during a given day?

P(between 5 and 7 silly t-shirts)
= P(5 silly t-shirts) + P(6 silly t-shirts) + P(7 silly t-shirts)
=(e^-17)*(17^5)/5! + (e^-17)*(17^6)/6! + (e^-17)*(17^7)/7!
=%v{ Math.exp(-17)*17**5/(5*4*3*2)+ Math.exp(-17)*17**6/(6*5*4*3*2)+ Math.exp(-17)*17**7/(7*6*5*4*3*2) } ≈ %v{(1000*(Math.exp(-17)*17**5/(5*4*3*2)+ Math.exp(-17)*17**6/(6*5*4*3*2)+ Math.exp(-17)*17**7/(7*6*5*4*3*2))).round()/1000.0}

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Assistment

Assistment #26859

You are previewing content.

In a particular county, the average number of suicides reported each month is %v{a}. Assume that the number of suicides follows a Poisson distribution.

C) Rounding to three digits after the decimal point, What is the probability that six or more suicides will be reported?

Comment on this question

Worked Example:

Suppose the number of silly t-shirts being reported at MIT during a given day follows a Poisson distribution with mean 17. What is the probability that more than three silly t-shirts are reported during a given day?

P(more than 3 silly t-shirts) = 1 - P(less than 3 silly t-shirts)
=1 - (P(0 silly t-shirts) + P(1 silly t-shirt) + P(2 silly t-shirts))
=1 - ((e^-17)*(17^0)/0! + (e^-17)*(17^1)/1! + (e^-17)*(17^2)/2!)
=1 - %v{ Math.exp(-17)*17**3/(3*2)+ Math.exp(-17)*17**2/2+ Math.exp(-17)*17**1 }
=%v{1- Math.exp(-17)*17**3/(3*2)-Math.exp(-17)*17**2/2- Math.exp(-17)*17**1 }
≈ %v{(1000*(1-Math.exp(-17)*17**3/(3*2)- Math.exp(-17)*17**2/2- Math.exp(-17)*17**1)).round ()/1000.0}

Comment on this hint

Type your answer below (mathematical expression):

-

Submit Answer

Appendix B
Appendix B

	Scaffolding	Pen-pencil
1	5	83
2	-17	-33
3	42	50
4	83	33
5	-40	-17
6	60	-33
7	33	33
8	-34	33
9	-33	17
10	34	-33
11	-17	-17
12	16	-33
13	-34	-17
14	-17	16
15	0	67
16	33	-17
17	0	50
18	16	0
19	19	16
20	16	0
21	0	17
22	16	17
23	16	-33
24	0	
25	0	
26	33	
27	0	
28	0	
29	-17	
30	-33	
31	50	
32	0	
Mean	8.25	8.65
SD	35.15	34.31

Appendix C

Scaffolding Problem Improvements View

Scaffolding Problem Accuracy	Post Test Accuracy	Accuracy Difference
69%	100%	50.00%
100%	33%	-67.00%
100%	100%	0.00%
100%	66%	-34.00%
76%	33%	-17.00%
100%	100%	0.00%
77%	100%	25.00%
100%	33%	-67.00%
53%	33%	-17.00%
100%	0%	-100.00%
69%	33%	-17.00%
55%	100%	25.00%
77%	33%	-42.00%
100%	100%	0.00%
100%	66%	-34.00%
77%	33%	-42.00%
100%	33%	-67.00%
91%	33%	-33.00%
100%	33%	-67.00%
100%	100%	0.00%
66%	66%	-9.00%
100%	100%	0.00%
88%	0%	-75.00%
75%	66%	-9.00%
62%	33%	-42.00%
100%	66%	-34.00%
100%	33%	-67.00%

Appendix C

100%	100%	0.00%
100%	66%	-34.00%
76%	66%	16.00%
100%	33%	-67.00%
75%	66%	-9.00%
100%	100%	0.00%
100%	100%	0.00%
88%	33%	-42.00%
100%	66%	-34.00%
100%	66%	-34.00%
100%	66%	-34.00%
66%	0%	-75.00%
75%	66%	-9.00%
100%	0%	-100.00%
100%	33%	-67.00%
100%	33%	-67.00%
53%	100%	50.00%
62%	33%	-42.00%
77%	33%	-42.00%
100%	0%	-100.00%
100%	66%	-34.00%
100%	66%	-34.00%
100%	66%	-34.00%
100%	100%	0.00%
100%	66%	-34.00%
100%	33%	-67.00%
56%	66%	16.00%
100%	66%	-34.00%
62%	33%	-42.00%

75%	33%	-42.00%
100%	100%	0.00%

Pen-paper Group Improvements View

Pen-Paper Group Pre-Test Accuracy	Pen-Paper Post Test Accuracy	Accuracy Difference
75%	100%	25.00%
100%	66%	-34.00%
50%	0%	-50.00%
100%	100%	0.00%
100%	100%	0.00%
50%	100%	50.00%
75%	33%	-42.00%
100%	100%	0.00%
25%	33%	8.00%
100%	66%	-34.00%
50%	100%	50.00%
100%	66%	-34.00%
100%	66%	-34.00%
75%	100%	25.00%
100%	0%	-100.00%
75%	66%	-9.00%
100%	33%	-67.00%
75%	0%	-75.00%
100%	100%	0.00%
25%	66%	41.00%
100%	66%	-34.00%
75%	66%	-9.00%
50%	33%	-17.00%

Appendix C

100%	100%	0.00%
50%	66%	16.00%
100%	33%	-67.00%
75%	33%	-42.00%
100%	66%	-34.00%
100%	100%	0.00%

MA2611 **2007A**

Lab 2: Solution for Pre-test

2007.9.12

Ryung Kim

Teaching Assistants: Dayang Liu and Yiwen Li

You will be asked to solve another set of problems after you go through these solutions. These problems will be similar to the quiz problems on this Friday. If you are well aware of the contents in chapter 3, you may need very little time to complete both tests. Please leave the classroom quietly so that you don't disturb your classmates who are still learning.

Q1. A manufacturer of roofing shingles wants to compare the performance of shingles with two different types of backings in field tests. To do so, they randomly select 30 communities around the county. In each, they randomly select a single-family house among those volunteered by their owners in response to an ad for a "free roof." They randomly select half the houses to receive one type of shingle and roof the rest with the second type. Various measures of the condition of each roof are obtained over a period of years.

What type of study is this one?

Recall that treatment is the condition the investigator wants to compare and response variable measures the consequence due to different treatments. So here the **treatments are 2 types of shingles**, and **the response variables are measures of the condition of each roof**. Now note that these **treatments are randomly assigned to subjects by the investigator**: "They randomly select half the houses to receive one type of shingle and roof the rest with the second type."

Since the treatments were assigned randomly, we can decide that this study is a **controlled experiment**. Furthermore, this **study does not have a block design**. Recall in some studies, we divide the experimental units into subgroups (blocks) that are expected to have similar response to common treatment. That was not the case here.

So in conclusion, this study has a **Completely Randomized Design**.

Q2. Consider the previous problem: another manufacturer is interested in the comparison of two types of shingles. This new manufacturer suspects that the annual precipitation affects the condition of roofs. So he divides the county into three regions: high precipitation, moderate precipitation, and low precipitation area. Within each of three regions, the manufacturer randomly assigns one type of shingle to 5 houses and another type to another 5 houses.

What type of study is this one?

Treatment and response variable are same as the first question. And it is still a **controlled experiment**. Remember because there was randomization in the study, it cannot be an observational study. Here, **the investigator used blocking** to minimize the effects of different precipitation amount. Before the experiment, he divided the experimental units into subgroups (blocks) that are expected to have similar response to common treatment.

So in conclusion, this study has a **Randomized Completely Block Design**.

Q3. Consider the previous problem again: another manufacturer is interested in the comparison of two types of shingles. However, this manufacturer does not have time or finance to assign free roof. Instead, she found 100 houses with their roof life span longer than 10 years, and also found 40 houses with their roof life span less than 3 years. Within the long-lasting roofs, 80% had type A shingles and within short-lived roofs, 30% had type B shingles.

What type of study is this one?

Treatment and response variable are same as the first question. And this is an observational study because there was no randomization.

Note that, at the outset, **houses were collected by their responses (roof condition)** not by the treatments (shingle types) they received: "Instead, she found 100 houses with their roof life span longer than 10 years, and also found 40 houses with their roof life span less than 3 years." Afterward, she studied what was the treatment (shingle types) of each house. And so we can conclude that this is a **retrospective observational study**.

Remember the following: 1) Prospective observational study establishes 'treatment' and 'control' groups at the outset and follows to observe the response. 2) Retrospective observational study first observes the end result (e.g. long or short life span), and differences in the treatments (e.g. shingle types) are sought.

Q4. Consider the previous problem again: another manufacturer first identified 1000 houses with shingle type A and 500 houses with shingle type B in the county. Among the 1000 houses, 20% had life span shorter than 3 years, and among 500 houses, 50% had life span shorter than 3 years.

What type of study is this one?

By now you should know that this is also an observational study. Note that, at the outset, the houses were collected by the treatments (shingle types) they received, not by their responses (roof condition) . And so, this is a **prospective observational study**.

In addition, let's note that **we cannot conclude that shingle type was the cause** of different life span of roofs. This is because we can only make such causal conclusion when the investigator randomly assigned the treatments. That is, only in controlled experiment, you can make such conclusion. Here, again, this is an observational study."

MA2611 **2007A**

Lab 3: Solution for Pre-test

2007.9.19

Ryung Kim
Teaching Assistants: Dayang Liu and Yiwen Li

You will be asked to solve another set of problems after you go through these solutions. These problems will be similar to the quiz problems on this Friday. If you are well aware of the contents in chapter 3, you may need very little time to complete both tests. Please leave the classroom quietly so that you don't disturb your classmates who are still learning.

Review Q1: Computing normal probability.
What is the probability that a normal random variable with mean 100 and standard deviation 25 to be greater than 125? You need to use Z-transformation and use R command pnorm().

P(X>125) = P(Z > (125-100)/25) = 1- P(Z < (125-100)/25) =0.1587
We can compute this in R by typing in 1- pnorm((125-100)/25)

Review Q2: Standard Deviation of sample mean
Each of x1, x2, x3,...,xn has standard deviation δ. Then, what's the standard deviation of the sample mean?

The standard deviation of the sample mean is δ/sqrt(n).

Q1. In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 4 random infants to be greater than 3000 grams?

THE DISTRIBUTION OF SAMPLE MEAN IS UNKNOWN BECAUSE ORIGINAL DISTRIBUTION OF ONE INFANT IS UNKNOWN. UNLESS WE HAVE LARGE ENOUGH SAMPLES TO USE CENTRAL LIMIT THEOREM, WE CANNOT COMPUTE PROBABILITY.

So in conclusion, we do not have enough information to answer this question.

Q2. In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is unknown. What is the probability that the mean weights of 40 random infants to be greater than 3400 grams?

THE DISTRIBUTION OF SAMPLE MEAN IS NORMAL BECAUSE WE HAVE LARGE ENOUGH SAMPLE SIZE EVEN THOUGH THE ORIGINAL DISTRIBUTION OF ONE INFANT IS UNKNOWN. As the mean of the birth weight of each infant is 3500 grams, **the mean of the sample mean of 40 infants' birth weights is also 3500 grams**. Recall that the mean of the sample mean is identical to the mean of each random variable. Here, each random variable (birth weight of each infant) has mean 3500 grams. As the birth weight of any infants has standard deviation of birth weight at 430 grams, **the standard deviation of the sample mean of 40 infants' birth weights is 67.99 grams.** We just used the property in the Review Q2 by computing δ/sqrt(n) with δ=430 grams, and n=25.

Now, we know that the sample mean has 1) normal distribution, 2) mean 3500, and 3) standard deviation 67.99. So the problem turns into following regular normal probability computation:

P(sample mean > 3400) = 1- P(sample mean < 3400) = 1- P(Z < (3400-3500)/67.99).

This can be solved in R by the following command: 1-pnorm((3400-3500)/67.99) .
In conclusion, the probability is 0.929.

Q3. In Norway, birth weights for infants whose gestational age is 40 weeks have mean 3500 grams and standard deviation 430 grams. Assume that the birth weight distribution is normal. What is the probability that the mean weights of 4 random infants to be greater than 3000 grams?

THE DISTRIBUTION OF SAMPLE MEAN WEIGHT IS NORMAL BECAUSE THE DISTRIBUTION OF EACH INFANT'S WEIGHT IS NORMAL. As the mean birth weight of each infant is 3500 grams, **the mean of the sample mean of 4 random infants' birth weights is 3500 grams**. Again, this is because the mean of the sample mean is identical to the mean of each random variable. Here, each random variable (birth weight of each infant) has mean 3500 grams. As the birth weight of any infants has standard deviation of birth weight 430 grams, **the standard deviation of the sample mean of 4 infants' birth weights is 215 grams**. Again, we just used the property in Review: Q2 and plug in the number into the formula: δ/sqrt(n). Here, δ=430 grams, and n=4.

Now we know that the sample mean has 1) normal distribution, 2) mean 3500, and 3) standard deviation 215. So the problem turns into following regular normal probability computation:

P(sample mean > 3000) = 1- P(sample mean < 3000) = 1- P(Z < (3000-3500)/215)

This can be solved in R by the following command: 1-pnorm((3000-3500)/215). So in conclusion, the probability is 0.99

Q4. The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery.
Assume that the battery life has a normal distribution. Suppose there are 5 batteries on hand. What is the probability that the 5 batteries are used up in less than 4000 hours?

Since total hours of 4000 is equivalent to mean of 800 hours, let's restate this problem in terms of sample mean of battery life, not of the total battery life: "What is the probability that the mean life of 5 batteries are less than 800 hours?"
Now let's solve the problem after we converted. **THE DISTRIBUTION OF SAMPLE MEAN IS NORMAL** BECAUSE THE ORIGINAL DISTRIBUTION OF ONE BATTERY IS NORMAL. Since the mean life time of each battery is 800 hours, **the mean of sample mean of randomly chosen 5 batteries is 800 hrs**. Again, this is because the mean of the sample mean is identical to the mean of each random variable. Here, each

random variable (life of each battery) has mean 800 hrs. Since the life of each battery has standard deviation of 150 hours, **the standard deviation of the sample mean of 5 batteries is 67.08 hrs**. We just used the property Review: Q2 and plug in the number into the formula: δ/sqrt(n). Here, δ=150 hours, and n=4.
Now we know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 67.08. So the problem turns into the following regular normal probability computation:

P(total hours < 4000) = P(sample mean < 8000) = P(Z < (800-800)/67.08)

This can be solved in R by the following command: pnorm((800-800)/67.08). So in conclusion, the probability is 0.5

Q5. The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery. Assume that the distribution of battery life is unknown. What is the probability that the 5 batteries are used up in less than 4000 hours?

Since total hours of 4000 is equivalent to mean of 800 hours, let's restate this problem in terms of sample mean of battery life, not of the total battery life: "What is the probability that the mean life of 5 batteries are less than 800 hours?" **THE DISTRIBUTION OF SAMPLE MEAN IS UNKNOWN** BECAUSE ORIGINAL DISTRIBUTION OF ONE BATTERY LIFE IS UNKNOWN. UNLESS WE HAVE LARGE ENOUGH SAMPLES TO USE CENTRAL LIMIT THEOREM, WE CANNOT COMPUTE PROBABILITY. In conclusion, we do not have enough information to solve this problem.

Q6. The life of a certain brand battery has mean 800 hours and a standard deviation of 150 hours. When one battery fails, it is immediately replaced by an identical new battery. Assume that the distribution of battery life is unknown. Suppose there are 30 batteries on hand. What is the probability that the 30 batteries are used up in less than 25000 hours?

Since total hours of 25000 is equivalent to mean of 833.33 hours, let's restate this problem in terms of sample mean of battery life, not of the total battery life: "What is the probability that the mean life of 30 batteries are less than 853.33 hours?"
THE DISTRIBUTION OF SAMPLE MEAN IS NORMAL BECAUSE WE HAVE LARGE ENOUGH SAMPLE SIZE EVEN THOUGH THE ORIGINAL DISTRIBUTION OF ONE INFANT IS UNKNOWN. Since the mean life time of each battery is 800 hours, **the mean of sample mean of randomly chosen 30 batteries is 800 hrs**. Again, this is because the mean of the sample mean is identical to the mean of each random variable. Here, each random variable (life of each battery) has mean 800 hrs. Since the life of each battery has standard deviation of 150 hours, **the standard deviation of the sample mean of 30 batteries is 27.39 hours**. We just used the property Review: Q2 and plug in the number into the formula: δ/sqrt(n). Here, δ=150 hours, and n=30.

Now we know that the sample mean has 1) normal distribution, 2) mean 800, and 3) standard deviation 27.39. So the problem turns into the following regular normal probability computation:

P(total hours < 25000) = P(sample mean < 833.33) = P(Z < (833.33-800)/27.39)

This can be solved in R by the following command: pnorm((833.33-800)/27.39). So in conclusion, the probability is 0.888