1. We calculate that, in order to estimate the true mean annual salary of all workers in a large union to within \$400 with 95% confidence, we need a sample of 500 workers. What sample size would be required to estimate the true mean salary to within \$500 with 95% confidence?

(A) 320 (B) 400 (C) 448 (D) 625 (E) 782

2. The sizes of farms in a U.S. state follow a normal distribution with standard deviation 30 acres. Suppose we measure the size of a random sample of farms and calculate the 98% confidence interval to be (295, 305). What is the correct interpretation of this interval?

   (A) Approximately 98% of farms have a size between 295 and 305 acres.

   (B) Approximately 98% of samples of 30 farms will have a mean size between 295 and 305 acres.

   (C) The probability that the population mean is between 295 and 305 acres is 98%.

   (D) In repeated samples of the same size, 98% of similarly constructed intervals will contain the sample mean.

   (E) In repeated samples of the same size, 98% of similarly constructed intervals will contain the population mean.

3. We would like to conduct a hypothesis test to examine whether there is evidence that the true mean amount spent on textbooks by a U of M student in one semester differs from \$400. A random sample of 50 students is selected and the mean amount they spent on textbooks for one semester is calculated to be \$430. Assume the population standard deviation is known to be \$165. What is the P-value for the appropriate hypothesis test?

(A) 0.1970 (B) 0.1112 (C) 0.0630 (D) 0.0985 (E) 0.1260

4. We take a simple random sample of 16 adults and ask them how long they sleep on a typical night. The sample mean is calculated to be 7.2 hours. Suppose it is known that the number of hours adults sleep at night follows a normal distribution with standard deviation 1.8 hours. An 88% confidence interval for the true mean time adults sleep at night is:

(A) (6.80, 7.60)

(B) (6.30, 8.10)

(C) (6.60, 7.80)

(D) (6.70, 7.70)

(E) (6.50, 7.90)

5. Which of the following statements comparing the standard normal distribution and the t distributions is **false**?

(A) The density curve for Z is taller at the center than the density curve for T.

(B) The t distributions have more area in the tails than the standard normal distribution.

(C) In tests of significance for μ, Z should be used as the test statistic when the distribution of X is normal, and T should be used in other cases.

(D) As the sample size increases, the t distribution approaches the standard normal distribution.

(E) In tests of significance for μ, T should be used as the test statistic only when the population standard deviation is unknown.

6. The Manitoba Department of Agriculture would like to estimate the average yield per acre of a new variety of corn for farms in southwestern Manitoba. It is desired that the final estimate be within five bushels per acre of the true mean yield. Due to cost restraints, a sample of no more than 65 one-acre plots of land can be obtained to conduct an experiment. The population standard deviation is known to be 17.33 bushels per acre. What is (approximately) the maximum confidence level that could be attained for a confidence interval that meets the Agriculture Department's specification?

(A) 90% (B) 95% (C) 96% (D) 98% (E) 99%

7. A statistical test of significance is designed to:

(A) prove that the null hypothesis is true.
(B) prove that the alternative hypothesis is true.
(C) find the probability that the null hypothesis is true.
(D) assess the strength of evidence in favour of the null hypothesis.
(E) assess the strength of evidence in favour of the alternative hypothesis.

8. We take a random sample of 12 observations from a normally distributed population. These 12 observations have a mean of 71.3 and standard deviation of 5.9. A confidence interval for μ is calculated to be **(67.335, 75.265)**. The confidence level of this interval is closest to:

(A) 90% (B) 95% (C) 96% (D) 98% (E) 99%

The next **two** questions (**9** and **10**) refer to the following:

The systems department of the Texaco Oil Company runs a large number of simulation programs on their mainframe computer. A manufacturer of new simulation software claims their program will run the simulation faster than the current mean of 30 minutes. The new program will be run 20 times. A hypothesis test of $H_0 : \mu = 30$ vs. $H_a : \mu < 30$ is to be conducted, and it is decided that H_0 will be rejected if $\bar{x} \leq 28.3$ minutes. Run times for the new program are known to follow a normal distribution with standard deviation 3.7 minutes.

9. The level of significance of the test is closest to:

(A) 0.01 (B) 0.02 (C) 0.03 (D) 0.04 (E) 0.05

10. What is the probability of making a Type II error if the true mean run time for the new program is actually 27 minutes?

(A) 0.0192 (B) 0.0244 (C) 0.0384 (D) 0.0475 (E) 0.0582

The next **two** questions (**11** and **12**) refer to the following:

The weights of apples in a large orchard are known to follow a normal distribution with a standard deviation of 12.2 grams. A random sample of 15 apples is selected from the orchard. We would like to conduct a hypothesis test at the 5% level of significance to determine whether the true mean weight of all apples in the orchard is greater than 150 grams.

11. What is the power of the hypothesis test if the true mean weight of all apples in the orchard is actually 160 grams?

(A) 0.9115 (B) 0.9370 (C) 0.9500 (D) 0.9726 (E) 0.9918

12. In the previous question, all else remaining the same, which of the following would have resulted in a higher power?

(I) selecting a random sample of 25 apples
(II) if the true mean weight of all apples was actually 170 grams
(III) using a 10% level of significance

(A) I only

(B) I and II only

(C) I and III only

(D) II and III only

(E) I, II and III

13. A statistician conducted a test of $H_0 : \mu = 100$ vs. $H_a : \mu \neq 100$ for the mean μ of some normally distributed population. Based on the gathered data, the statistician calculated a sample mean of $\bar{x} = 110$ and concluded that H_0 could be rejected at the 5% level of significance. Using the same data, which of the following statements **must** be true?

(I) A test of $H_0 : \mu = 100$ vs. $H_a : \mu > 100$ at the 5% level of significance would also lead to rejecting H_0.

(II) A test of $H_0 : \mu = 90$ vs. $H_a : \mu \neq 90$ at the 5% level of significance would also lead to rejecting H_0.

(III) A test of $H_0 : \mu = 100$ vs. $H_a : \mu \neq 100$ at the 1% level of significance would also lead to rejecting H_0.

(A) I only

(B) II only

(C) I and II only

(D) I and III only

(E) I, II and III

14. We take random samples of individuals from two normally distributed populations and record the value of some random varialbe X. Some summary statistics are shown below:

Population	Sample Size	Sample Mean	Sample Std. Dev.
1	7	29	4.0
2	14	25	9.2

We would like to contstruct a 95% confidence interval for the difference $\mu_1 - \mu_2$ in population means. What is the value of the standard error of $\bar{X}_1 - \bar{X}_2$?

(A) 2.24 (B) 1.86 (C) 3.27 (D) 2.89 (E) 3.45

The next **three** questions (**15**, **16** and **17**) refer to the following:

Coke and Pepsi are the two most popular brands of cola on the market. Do consumers prefer either one of the two brands of cola over the other? We conduct an experiment as follows: 20 volunteers participate in a blind taste test. Each volunteer tastes both Coke and Pepsi (in random order) and scores the taste of each cola on a scale from 0 to 100. Some information that may be helpful is shown in the table below:

Scores for Coke	Scores for Pepsi	Difference (Coke − Pepsi)
mean = 78	mean = 83	mean = −5
std. dev. = 27	std. dev. = 24	std. dev. = 13

15. Which of the following statements is/are **true**?

(I) The scores for Coke and Pepsi for each individual are independent.
(II) The scores for Coke and Pepsi for each individual are dependent.
(III) In order to conduct the matched pairs t test, we must assume that scores for Coke and scores for Pepsi both follow normal distributions.
(IV) In order to conduct the matched pairs t test, we must assume that the difference in scores (Coke − Pepsi) follow a normal distribution.

(A) I only
(B) I and III only
(C) I and IV only
(D) II and III only
(E) II and IV only

16. We will use the critical value method to conduct the test, with a 10% level of significance. Assuming the appropriate assumptions are satisfied, the rejection rule is to reject H_0 if:

(A) $|t| \geq 1.328$ (B) $t \geq 1.645$ (C) $|t| \geq 1.729$ (D) $t \geq 1.282$ (E) $|t| \geq 1.833$

17. Assuming the appropriate assumptions are satisfied, what is the value of the test statistic for the appropriate test of significance?

(A) -0.38 (B) -7.69 (C) -2.85 (D) -1.72 (E) -0.88

18. We record the amount of time (in hours) spent watching TV per week for random samples of young children and teenagers. Some summary statistics are shown below:

	Sample Size	Mean	Standard Deviation
Children	16	26.4	7.3
Teenagers	9	19.5	6.2

Weekly TV viewing times follow normal distributions for both children and teenagers. An appropriate hypothesis test is conducted at the 1% level of significance to determine whether there is a difference in the true mean TV viewing time for children and teenagers. The test statistic is calculated to be $t = 2.39$. We conclude that we should:

(A) reject H_0, since the P-value is between 0.01 and 0.02.

(B) fail to reject H_0, since the P-value is between 0.01 and 0.02.

(C) reject H_0, since the P-value is between 0.02 and 0.04.

(D) fail to reject H_0, since the P-value is between 0.02 and 0.04.

(E) reject H_0, since the P-value is between 0.04 and 0.05.

19. Random samples of contestants who have appeared on two popular game shows (*Wheel of Fortune* and *Jeopardy*) are selected. The Wheel of Fortune contestants in the sample won an average of $5700 more than the Jeopardy contestants in the sample. A 95% confidence interval for the difference in the mean amount of money won by contestants on the two game shows is calculated to be (−1900, 13300). We would like to conduct a hypothesis test to determine whether the mean amounts won on the two game shows differ. Assume the appropriate normality assumptions are satisfied. We would:

(A) reject H_0 at the 5% level of significance since the value 0 is contained in the 95% confidence interval.

(B) fail to reject H_0 at the 5% level of significance since the value 0 is contained in the 95% confidence interval.

(C) fail to reject H_0 at the 5% level of significance since the value 5700 is contained in the 95% confidence interval.

(D) reject H_0 at the 10% level of significance since the value 5700 is contained in the 95% confidence interval.

(E) fail to reject H_0 at the 10% level of significance since the value 0 is contained in the 95% confidence interval.

The next **two** questions (**20** and **21**) refer to the following:

Several workers in a large office building suspect that coffee vending machine A dispenses more coffee on average than vending machine B. To test this belief, several samples were taken from each machine. The amounts (in ounces) dispensed and some summary statistics are given below:

Machine	n	$\bar{x}$	s
A	12	11.74	0.56
B	8	11.18	0.79

20. Assuming fill volumes follow a normal distribution for both machines, the test statistic for the appropriate test of significance is:

(A) $$\frac{11.74 - 11.18}{\sqrt{0.46\left(\frac{1}{12} + \frac{1}{8}\right)}}$$

(B) $$\frac{11.74 - 11.18}{\sqrt{\frac{(0.56)^2}{12} + \frac{(0.79)^2}{8}}}$$

(C) $$\frac{11.74 - 11.18}{\sqrt{0.43\left(\frac{1}{12} + \frac{1}{8}\right)}}$$

(D) $$\frac{11.74 - 11.18}{\sqrt{\frac{(0.56)^2}{11} + \frac{(0.79)^2}{7}}}$$

(E) $$\frac{11.74 - 11.18}{\sqrt{0.68\left(\frac{1}{12} + \frac{1}{8}\right)}}$$

21. We would make a Type II error in this test if we concluded that:

(A) Machine A dispenses more on average than Machine B when in fact Machine B dispenses more on average than Machine A.

(B) there is no evidence that Machine A and B dispense the same amount on average when in fact Machine A does dispense more on average.

(C) Machine B dispenses more on average than Machine A when in fact Machine A dispenses more on average than Machine B.

(D) Machine A dispenses more on average than Machine B when in fact Machines A and B dispense the same amount on average.

(E) there is no evidence that Machine A dispenses more on average than Machine B when in fact Machine A does dispense more on average.

22. We would like to conduct a hypothesis test at the 5% level of significance to determine whether hockey players weigh more on average than baseball players. We record the weights of a random sample of 12 professional hockey players and 7 professional baseball players. Weights of athletes in both sports are known to follow normal distributions. Some JMP output that may be helpful is shown below:

Means and Std Deviations

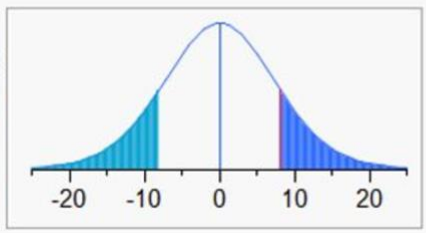
Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
Baseball	7	187.571	7.7213	2.9184	180.43	194.71
Hockey	12	195.583	18.3920	5.3093	183.90	207.27

t Test

Hockey-Baseball

Assuming equal variances

Difference	8.012	t Ratio	1.087592
Std Err Dif	7.367	DF	17
Upper CL Dif	23.554	Prob > |t|	0.2920
Lower CL Dif	-7.530	Prob > t	0.1460
Confidence	0.95	Prob < t	0.8540

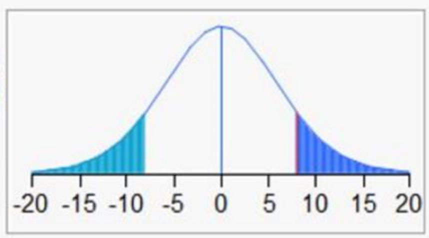


t Test

Hockey-Baseball

Assuming unequal variances

Difference	8.012	t Ratio	1.322418
Std Err Dif	6.059	DF	15.97728
Upper CL Dif	20.857	Prob > |t|	0.2046
Lower CL Dif	-4.833	Prob > t	0.1023
Confidence	0.95	Prob < t	0.8977



Based on the data, we would fail to reject the null hypothesis of the appropriate test of significance, since

(A) $t = 1.32 < t^* = 1.753$

(B) $t = 1.09 < t^* = 1.740$

(C) $t = 1.32 < t^* = 1.645$

(D) $t = 1.09 < t^* = 2.110$

(E) $t = 1.32 < t^* = 2.131$

23. We would like to conduct a hypothesis test to determine whether females (F) can complete a certain task faster than males (M) on average. We will commit a Type I Error if we conclude that:

(A) $\mu_F > \mu_M$ when in fact $\mu_F = \mu_M$

(B) $\mu_F < \mu_M$ when in fact $\mu_F > \mu_M$

(C) $\mu_F = \mu_M$ when in fact $\mu_F < \mu_M$

(D) $\mu_F < \mu_M$ when in fact $\mu_F = \mu_M$

(E) $\mu_F > \mu_M$ when in fact $\mu_F < \mu_M$

The next **two** questions (**24** and **25**) refer to the following:

We conduct an experiment to compare the responses to six different treatments. Volunteers are randomly assigned to receive one of the six treatments. The response variable is measured for each individual and an analysis of variance F test is conducted to assess the equality of the means for the six treatments. Responses for each of the six treatments are known to follow normal distributions. Some summary statistics are shown below:

Treatment	Sample Size	Sample Mean	Sample Std. Dev.
1	3	17	8
2	5	14	7
3	2	21	11
4	4	20	6
5	3	18	9
6	4	12	10

24. Under the null hypothesis, we assume all population means are equal. What is the estimate of this common population mean?

(A) 15.99 (B) 16.43 (C) 16.78 (D) 17.00 (E) 17.22

25. What is the 5% critical value for the appropriate test of significance?

(A) 2.57 (B) 2.68 (C) 2.79 (D) 2.85 (E) 2.90

The next **three** questions (**26** to **28**) refer to the following:

Battery life of MP3 players is of great concern to customers. A consumer group tested four brands of MP3 players to determine the battery life. Samples of players of each brand were fully charged and left to run on medium volume until the battery died. The following table displays the number of hours each of the batteries lasted:

Brand

	A	B	C	D
	24	26	18	27
	19	28	20	24
	22	24	21	22
		25		24
mean	21.67	25.75	19.67	24.25
std. dev.	2.52	1.71	1.53	2.06

We would like to conduct an analysis of variance to compare the performance of the four brands. Suppose that all necessary assumptions have been satisfied. The ANOVA table (with some values missing) is shown below:

Source of Variation	*df*	Sum of Squares	Mean Squares	*F*
Groups				
Error			3.88	
Total		113.71		

26. What is the alternative hypothesis for the appropriate test of significance?

(A) H_a: All population variances are different.

(B) H_a: At least one sample mean differs.

(C) H_a: All population means are different.

(D) H_a: At least one population variance differs.

(E) H_a: At least one population mean differs.

27. What is the P-value of the appropriate test of significance?

(A) less than 0.001

(B) between 0.01 and 0.025

(C) between 0.025 and 0.05

(D) between 0.05 and 0.10

(E) greater than 0.10

28. We would like to construct a 95% confidence interval for the true mean battery life for all Brand B MP3 players. The 95% confidence interval is:

(A) (23.56, 27.94)
(B) (23.17, 28.33)
(C) (23.03, 28.47)
(D) (22.51, 28.99)
(E) (22.04, 29.46)

The next **two** questions (**29** and **30**) refer to the following:

We conduct an experiment to compare the responses to five different treatments. Volunteers are randomly assigned to receive one of the five treatments. Each treatment is assigned to the same number of individuals. The response variable is measured for each individual and an analysis of variance F test is conducted to assess the equality of the means for the five treatments. Responses for each of the five treatments are known to follow normal distributions with common standard deviation. The ANOVA table (with some values missing) is shown below:

Source of Variation	*df*	Sum of Squares	Mean Squares	F
Groups		74.52		
Error	60		5.78	
Total				

29. How many individuals received each of the five treatments?

(A) 11 (B) 12 (C) 13 (D) 14 (E) 15

30. One assumption required in conducting and ANOVA F test is that all population standard deviations are equal. The estimate of this common population standard deviation is:

(A) 2.40 (B) 4.32 (C) 5.78 (D) 18.63 (E) 33.41