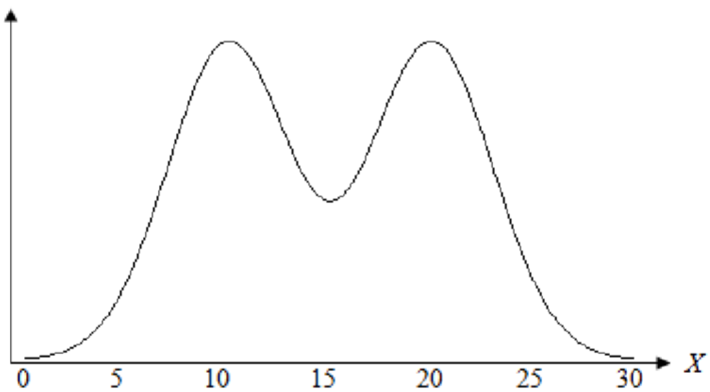
1. A bimodal probability distribution is one with two distinct peaks. A random variable X follows a bimodal distribution with mean 15 and standard deviation 4, as shown below:



We will take a random sample of 10,000 individuals from this distribution and calculate the sample mean $\bar{x}$. The sampling distribution of $\bar{X}$ is:

(A) approximately normal with mean close to 15 and standard deviation close to 0.0004.

(B) bimodal with mean close to 15 and standard deviation close to 0.04.

(C) approximately normal with mean close to 15 and standard deviation close to 0.04.

(D) bimodal with mean close to 15 and standard deviation close to 4.

(E) approximately normal with mean close to 15 and standard deviation close to 4.

2. GPAs at a large university follow a normal distribution with mean 2.84 and standard deviation 0.46. What is the probability that a random sample of four students has a mean GPA greater than 3.00?

   (A) 0.2420

   (B) 0.6957

   (C) 0.7580

   (D) 0.3043

   (E) impossible to calculate with the information given.

3. Annual salaries of workers in a large union follow a normal distribution with standard deviation \$10,000. What sample size is required if we want to estimate the true mean salary to within \$2,000 with 93% confidence?

(A) 82 (B) 85 (C) 88 (D) 94 (E) 97

4. We will take a random sample of 30 vehicles of a certain make and model and measure the fuel efficiency in miles per gallon (mpg) of each of them. We will conduct a hypothesis test at the 10% level of significance to determine whether there is evidence that the true mean fuel efficiency of all cars of this make and model differs from 32 mpg. What is the probability of failing to reject H_0 if the true mean is in fact 32 mpg?

   (A) 0.10 (B) 0.95 (C) 0.05 (D) 0.90 (E) 0.20

5. For which of the following tests would a Type II Error have more serious consequences than a Type I Error?

(I) A police officer is trying to decide whether to pull over a driver.
H_0: The driver is not drunk.
H_a: The driver is drunk.

(II) You are deciding whether to go for a swim in the ocean.
H_0: There are no sharks in the water.
H_a: There are sharks in the water.

(III) A shoplifter is deciding whether to steal an expensive camera from Walmart.
H_0: He will not be caught.
H_a: He will be caught.

(A) II only

(B) I and II only

(C) I and III only

(D) II and III only

(E) I, II and III

6. A horticulturist wishes to estimate the true mean growth of seedlings in a large timber plot last year. A random sample of 20 seedlings is selected and the one-year growth for each of them is measured. The sampled seedlings have a mean growth of 5.6 cm and a standard deviation of 1.5 cm. One-year seedling growth is known to follow a normal distribution with standard deviation 2.0 cm. We wish to conduct a hypothesis test at the 10% level of significance to determine whether the true mean one-year seedling growth differs from 6.0 cm. The critical region for the test is:

(A) $|z| > 1.645$

(B) $t < -1.328$

(C) $z < -1.282$

(D) $|t| > 1.729$

(E) $|z| > 1.282$

7. A random variable X follows a normal distribution with standard deviation $\sigma = 8$. We take a random sample of 25 individuals from the population and we calculate a confidence interval for μ to be **(36.7136, 43.2864)**. What is the confidence level of this interval?

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

8. A test of $H_0 : \mu = 100$ vs. $H_a : \mu \neq 100$ is conducted for the mean μ of some population. We take a sample of ten individuals and calculate a sample mean of 104. A 98% confidence interval for μ is calculated to be (101, 107). Which of the following statements is **true**?

   (A) We fail to reject H_0 at a 4% level of significance, since 100 is not contained in the 98% confidence interval.

   (B) We fail to reject H_0 at a 2% level of significance, since 104 is contained within the 98% confidence interval.

   (C) We reject H_0 at the 1% level of significance, since 100 is not contained within the 98% confidence interval.

   (D) We fail to reject H_0 at the 4% level of significance, since 104 is contained within the 98% confidence interval.

   (E) We reject H_0 at a 2% level of significance, since 100 is not contained within the 98% confidence interval.

9. A simple random sample of six male patients over the age of 65 is being used in a blood pressure study. The standard error of the mean blood pressure of these six men was calculated to be 7.8. What is the standard deviation of these six blood pressure measurements?

(A) 2.8 (B) 3.2 (C) 14.4 (D) 19.1 (E) 46.8

10. A statistician conducted a test of $H_0 : \mu = 1$ vs. $H_a : \mu > 1$ for the mean μ of some population. Based on the gathered data, the statistician concluded that H_0 could be rejected at the 1% level of significance. Using the same data, which of the following statements must be true?

(I) A test of $H_0 : \mu = 1$ vs. $H_a : \mu > 1$ at the 10% level of significance would also lead to rejecting H_0.

(II) A test of $H_0 : \mu = 0$ vs. $H_a : \mu > 0$ at the 1% level of significance would also lead to rejecting H_0.

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

(A) I only

(B) I and II only

(C) I and III only

(D) II and III only

(E) I, II and III

11. The Winnipeg Transit Commission claims that the average time taken by the Number 60 bus to travel from the University of Manitoba to downtown is 27 minutes. A student who takes this route often believes that the true mean time is greater than 27 minutes. The student records the time for a sample of five trips. These trips had a mean time of 30 minutes and a standard deviation of 4 minutes. Suppose it is known that trip times have a normal distribution. The P-value for the appropriate test of significance to test the student's suspicion is:

(A) between 0.01 and 0.02.

(B) between 0.02 and 0.025.

(C) between 0.025 and 0.05.

(D) between 0.05 and 0.10.

(E) between 0.10 and 0.15.

12. We conduct a hypothesis test of $H_0 : \mu = 20$ vs. $H_a : \mu \neq 20$ for the mean of some population at the 5% level of significance. Using a sample of size 30, we determine that H_0 should be rejected if $\bar{x} < 17$ or $\bar{x} > 23$. The power of the test against the alternative $H_a : \mu = 25$ is calculated to be 0.8621.

Which of the following statements is **true**?

(A) If we had instead used a sample size of 50, the probability of a Type II error would increase.

(B) If we used a 5% level of significance and a sample of size 50, the power of the test would decrease.

(C) The power of the test against the alternative $H_a : \mu = 15$ is 0.1379.

(D) If we had instead used a 1% level of significance, the probability of a Type II error would decrease.

(E) If we had instead used a 10% level of significance, the power of the test would increase.

13. Bottles of a certain brand of apple juice are supposed to contain 300 ml of juice. There is some variation from bottle to bottle because of imprecisions in the filling machinery. A consumer advocate inspector selects a random sample of 36 bottles and tests the hypotheses $\mathrm{H_0} : \mu = 300$ vs. $\mathrm{H_a} : \mu < 300$ at the 10% level of significance. Fill volumes are known to follow a normal distribution with standard deviation 3 ml. What is the power of the test if the true mean fill volume is actually 298 ml?

(A) 0.9925 (B) 0.9943 (C) 0.9957 (D) 0.9967 (E) 0.9977

The next **two** questions (**14** and **15**) refer to the following:

Eight males and eight females volunteered to be part of an experiment. All 16 people were Caucasian and between the ages of 18 and 24. Each of the eight male participants was randomly assigned a night club, and each of the eight females was randomly assigned to one of these same eight night clubs. One Friday night, all 16 people went out to the bar. Each person then went up to the bar to order a drink (within 5 minutes of one another, with the order randomly determined.) The time (in seconds) until each customer was served was recorded. The table below contains some information that may be helpful:

Females	Males	Difference (d = Female − Male)
mean = 48	mean = 62	mean = −14
std. dev. = 18	std. dev. = 23	std. dev. = 15

The question of interest in this experiment is whether females receive faster service at bars on average than males. Assume the appropriate normality assumptions are satisfied.

14. What are the hypotheses for the appropriate test of significance?

(A) $H_0 : \mu_d = 0$ vs. $H_a : \mu_d < 0$

(B) $H_0 : \bar{X}_d = 0$ vs. $H_a : \bar{X}_d > 0$

(C) $H_0 : \mu_F = \mu_M$ vs. $H_a : \mu_F > \mu_M$

(D) $H_0 : \bar{X}_d = 0$ vs. $H_a : \bar{X}_d < 0$

(E) $H_0 : \mu_d = 0$ vs. $H_a : \mu_d > 0$

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

(A) between 0.005 and 0.01
(B) between 0.01 and 0.02
(C) between 0.02 and 0.025
(D) between 0.025 and 0.05
(E) between 0.05 and 0.10

The next **two** questions (**16** and **17**) refer to the following:

The following data represent the fat content found in samples of two popular brands of ice cream:

Brand	Fat					sample mean	sample variance
A	6.3	5.1	6.8	6.9	7.4	6.50	0.77
B	6.3	5.7	5.9	6.4	5.1	5.88	0.27

16. We would like to conduct a hypothesis test to determine whether the true mean fat content differs for the two brands of ice cream. Assuming fat content follows a normal distribution for both brands, we should use:

(A) a matched pairs t test with 4 degrees of freedom.

(B) a pooled two-sample t test with 8 degrees of freedom.

(C) a conservative two-sample t test with 7 degrees of freedom.

(D) a pooled two-sample t test with 9 degrees of freedom.

(E) a conservative two-sample t test with 8 degrees of freedom.

17. The P-value for the appropriate test of significance is calculated to be 0.21. We interpret this value to mean:

(A) If the true mean fat content was equal for the two brands, the probability of incorrectly concluding that the means differ would be 0.21.

(B) The probability that the true mean fat content is equal for the two brands is 0.21.

(C) The probability that the true mean fat content differs for the two brands is 0.21.

(D) If the true mean fat content were equal for the two brands, the probaility of observing a difference in sample means at least as extreme as 0.62 would be 0.21.

(E) If the true mean fat content differed for the two brands, the probability of observing a difference in sample means at least as extreme as 0.62 would be 0.21.

18. The distances (in km) travelled by random samples of University of Manitoba and University of Winnipeg students each day (from home to their university) are recorded and some summary statistics are shown below:

	n	$\bar{x}$	s
U of M	11	9.6	3.2
U of W	7	5.5	2.4

Distances are known to follow normal distributions for students at both universities. We would like to conduct a hypothesis test to determine whether U of M students travel further on average than U of W students. The P-value for the appropriate test of significance is:

(A) between 0.01 and 0.02

(B) between 0.02 and 0.025

(C) between 0.025 and 0.05

(D) between 0.05 and 0.10

(E) between 0.10 and 0.15

19. We take independent simple random samples from two normally distributed populations. Some summary statistics are shown in the table below:

Sample	Sample Size	Sample Mean	Sample Std. Dev.
1	12	43	16
2	7	37	9

We would like to construct a 90% confidence interval for the difference in the two population means $\mu_1 - \mu_2$. The margin of error for the confidence interval is:

(A) 5.74 (B) 6.63 (C) 7.65 (D) 9.98 (E) 11.53

20. We would like to conduct a pooled two-sample t test to compare the means of two populations. Which of the following assumptions are necessary?

(I) Both populations are normally distributed.
(II) The two samples are dependent.
(III) The two samples are independent.
(IV) The population standard deviations are known.
(V) The population standard deviations are equal.
(VI) The sample standard deviations are equal.

(A) III, IV and V

(B) I, III, V and VI

(C) I, III and VI

(D) I, II, IV and VI

(E) I, III and V

21. In which of the following situations are the matched pairs t procedures appropriate?

(I) We would like to compare the mean number of mosquitoes caught in two different types of mosquito traps. A sample of 20 locations is selected in Winnipeg, and one of each type of trap is placed at each location.

(II) We would like to compare the mean commuting times from home to work for workers in Winnipeg and Toronto. A sample of 30 workers in each city is selected.

(III) We would like to compare the mean summer temperatures in Winnipeg and Brandon (Manitoba's second largest city, 200 km west of Winnipeg). On a sample of 25 days, we measure the temperature in each city.

(A) I only

(B) II only

(C) III only

(D) I and II only

(E) I and III only

22. The serum calcium measurements for samples of two breeds of cows are shown below. Serum calcium levels are known to follow a normal distribution for both breeds.

Chester White	116	112	82	63	117	69
Hampshire	62	79	80	85	60	71

We would like to conduct a hypothesis test to determine whether the true mean serum calcium level differs for the two breeds. 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%
Hampshire	6	72.8333	10.2258	4.175	62.102	83.56
Chester White	6	93.1667	24.7501	10.104	67.193	119.14

t Test

Chester White-Hampshire

Assuming equal variances

Difference	20.333	t Ratio	
Std Err Dif	10.933	DF	10
Upper CL Dif	44.693	Prob > |t|	
Lower CL Dif	-4.026	Prob > t	
Confidence	0.95	Prob < t	

t Test

Chester White-Hampshire

Assuming unequal variances

Difference	20.333	t Ratio	
Std Err Dif	10.933	DF	6.658692
Upper CL Dif	46.456	Prob > |t|	
Lower CL Dif	-5.789	Prob > t	
Confidence	0.95	Prob < t	

The P-value for the appropriate test of significance is:

(A) between 0.01 and 0.02

(B) between 0.025 and 0.05

(C) between 0.05 and 0.10

(D) between 0.10 and 0.20

(E) between 0.20 and 0.30

23. A student would like to conduct a hypothesis test to determine whether the true mean exam score for students in Dr. E. Zee's class is greater than that for student's in Dr. D. Ficult's class. Exam scores are known to follow a normal distribution for each of the two classes. The student selects random samples of students from each professor's class. Some summary statistics for their exam scores are shown below:

Group	Sample Size	Sample Mean	Sample Std. Dev.
Dr. E. Zee	14	74	6
Dr. D. Ficult	10	53	14

What is the value of the test statistic for the appropriate test of significance?

(A) 3.77 (B) 4.46 (C) 5.04 (D) 6.62 (E) 7.13

24. A horticulturist is investigating the phosphorous content of tree leaves from three different varieties of apple trees. Phosphorous content is known to follow a normal distribution for each of the three varieties. Random samples of four leaves from each of the three varieties are analyzed for phosphorous content. Population variances are known to be equal for all three varieties. We would like to test the hypothesis of equal population means for the three varieties at the 10% level of significance. The critical value for the appropriate hypothesis test is:

(A) 1.36 (B) 2.81 (C) 3.01 (D) 5.22 (E) 9.38

25. We take simple random samples from two normally distributed populations and measure the value of some variable X. We conduct a pooled two-sample t test to determine whether there is evidence that the population means differ. We calculate a test statistic of $t = 2.47$ and a P-value of 0.04. Suppose we had instead conducted the test using an analysis of variance. What would be the value of the test statistic and the P-value?

(A) 6.10 and 0.04

(B) 2.47 and 0.0016

(C) 1.57 and 0.04

(D) 6.10 and 0.0016

(E) 2.47 and 0.20

26. We are conducting an analysis of variance to compare the means of four populations. In conducting this test, we will commit a Type I error if we:

(A) fail to conclude that all four means differ when in fact this is the case.

(B) conclude that at least one of the means differs when in fact they are all equal.

(C) conclude that all four means differ when in fact they are all equal.

(D) fail to conclude that at least one mean differs when in fact this is the case.

(E) conclude that one mean differs when in fact all four means differ.

27. We would like to compare the means of three normally distributed populations with common variance. We conduct three separate pooled two-sample t tests and we obtain the P-values shown below:

Test 1	Test 2	Test 3
$\mathrm{H_0}$: $\mu_1 = \mu_2$	$\mathrm{H_0}$: $\mu_1 = \mu_3$	$\mathrm{H_0}$: $\mu_2 = \mu_3$
$\mathrm{H_a}$: $\mu_1 \neq \mu_2$	$\mathrm{H_a}$: $\mu_1 \neq \mu_3$	$\mathrm{H_a}$: $\mu_2 \neq \mu_3$
P-value = 0.02	P-value = 0.23	P-value = 0.11

Suppose we were to conduct an analysis of variance F test at the 5% level of significance to compare the three means. Which of the following statements is true?

(A) The P-value of the ANOVA F test would be 0.02 (the smallest of the P-values for the three two-sample tests).

(B) The P-value of the ANOVA F test would be 0.12 (the average of the P-values for the three two-sample tests).

(C) The null hypothesis would be rejected, since one of the three two-sample tests had a P-value less than 0.05.

(D) The null hypothesis would not be rejected, since two of the three two-sample tests had a P-value greater than 0.05.

(E) There is no way to determine the results of the ANOVA F test from the three two-sample tests. If we wanted to compare all three means, we should have conducted an analysis of variance test to begin with.

The next **three** questions (**28** to **30**) refer to the following:

A medical researcher wants to compare three different techniques designed to lower blood pressure in patients with hypertension. Five subjects are assigned to each of the three groups. The first group exercises, the second group meditates and the thrid group takes medication. After five weeks, the reduction in each person's blood pressure is recorded. The researcher wants to determine if there is any difference in the effectiveness of the three techniques. The ANOVA table (with some values missing) is shown below:

Source of Variation	*df*	Sum of Squares	Mean Squares	F
Groups		102		
Error				
Total		282		

28. What is the P-value of the test?

(A) less than 0.001

(B) between 0.001 and 0.01

(C) between 0.01 and 0.025

(D) between 0.025 and 0.05

(E) between 0.05 and 0.10

29. What is the estimate of the common variance for the three treatments?

(A) 4 (B) 7 (C) 13 (D) 15 (E) 51

30. We would like to construct a 95% confidence interval for the difference in the mean response for the exercise and medication groups. The margin of error for the appropriate confidence interval is:

(A) 5.25 (B) 5.34 (C) 5.65 (D) 9.53 (E) 12.49