

I confirm that, for this exam, I am allowed only a calculator, the provided probability tables, and two-sides of an 8.5x11" sheet of paper for notes. I will not look at anybody else's exam and I will take all necessary efforts to prevent others from seeing my exam. I will use the provided probability tables to cover my work and answers. The consequence of using additional test aids, copying from others, or allowing others to copy my work can result in an F grade for this class.

I have read and agree with the above statement. Signature: _____

(a.) Show all work to receive full credit.

(b.) Circle your final answer.

(c.) Ask the instructor for clarification if any questions are unclear.

(1) Suppose you are estimating the mean monthly rent in San Francisco for an apartment. Your random sample of $n=36$ apartments provided a sample mean of \$1200 and a sample standard deviation of \$180.

(1a. 4pts) Construct a 95% confidence interval for the mean monthly rent for an apartment in San Francisco. **Show your work.**

Answer= (_____, _____)

(1b. 3pts) Suppose your answer for (1a) is (L,U), where L and U are the lower and upper bounds for the 95% confidence interval. Circle the phrase which best fills the blank.

"You are 95% confident that the mean monthly rent of _____ is between L and U dollars."

(i) the 36 sampled apartments (ii) apartments in San Francisco

(1c. 2pts) Circle the word that best completes the sentence: Increasing the sample size to 90 apartments would most likely make the width of the 95% confidence interval _____.

(i) narrower (ii) unchanged (iii) wider.

(1d. 2pts) \$1200 is a _____. Circle: (i) parameter (ii) statistic

(2. 3pts) When performing a statistical test, we do our calculations assuming the _____ hypothesis is true. Circle the correct answer to fill in the blank: (i) null (ii) alternative (iii) chi-square (iv) random

(3. 2pts) Calculate the **standard deviation** for the three values: 2, 3, 10. (Circle answer and show work.)

(4. 4pts) Use the provided three options to describe A, B, C, and D in the below table about correct and incorrect decisions for hypothesis testing. **Answers may be used more than once.**

		Reality(truth)	
		Null hypothesis is true	Alternative hypothesis is true
Decision	Keep null hypothesis	A	B
	Reject null hypothesis	C	D

Options: "Correct conclusion", "Type 1 error", "Type 2 error"

A= _____ B= _____

C= _____ D= _____

(5. 3pts) Circle the statement below that best describes a p-value.

- The p-value is the probability of the null hypothesis being true.
- The p-value is the probability of the alternative hypothesis being true.
- The p-value is the probability of committing a type 1 error.
- The p-value is the probability of committing a type 2 error.
- Assuming the null hypothesis is true, the p-value is the probability of getting a test statistic as extreme or more extreme than the test statistic you calculated from your sampled data.
- Assuming the alternative hypothesis is true, the p-value is the probability of getting a test statistic as extreme or more extreme than the test statistic you calculated from your sampled data.

(6. 2pts) Suppose you are performing a statistical test using a level of significance of 5%; that is, $\alpha=0.05$. If the null hypothesis is true, what is the probability of sampling data such that the statistical test will reject the null hypothesis?

Answer: _____

(7. 3pts) Circle **True** or **False**: Suppose you are performing a statistical test using a level of significance of 5%; that is, $\alpha=0.05$. If your calculated p-value is 0.37, then you should reject the null hypothesis.

(8. 2pt) Use one sentence to describe what makes an experiment be a double-blind experiment?

(9. 2pt) Explain in one or two sentences, why the median rather than mean is typically used to describe housing prices.

(10.) A large car company colors cars according to customer orders, but the color must be either black, white, or blue. Nationally, 30% of the cars ordered are black, 20% white, and 50% blue. In Humboldt County, however, out of 200 cars ordered, 55 were black, 30 white, and 115 blue. In the following steps, you will perform a chi-square test on these data to test whether Humboldt County car color preference is statistically different from national car color preference. Use $\alpha=0.05$.

(10a. 2pts) Assuming Humboldt County's color preference is no different from national color preference, calculate the expected numbers of
black = _____, white = _____, and blue = _____ cars.
Show your work.

(10b. 2pts) Chi-square statistic = _____. Show your work.

(10c. 2pts) Determine the critical value: df = _____, critical value = _____.

(10d. 2pts) Circle **True** or **False**: The results of the statistical test suggest you should keep the null hypothesis that Humboldt County's color preferences are the same as the national color preferences. (Hint: Use your result from parts b and c.)

(11. 3pts) Circle: **True** or **False**. The Central Limit Theorem tells us that as the sample size increases, the distribution of the sampled values becomes more like the normal distribution.

(12. 3pts) The total area beneath the curve of a probability density function for a continuous random variable is _____.

(13.) A company claims *on average* there are 12 ounces of cereal in their boxes of cereal. A suspicious statistician performed a 1-sample t- test was to determine if the mean weight of the company's cereal boxes was less than 12 ounces. The contents from 36 boxes were measured resulting in a mean of 11.85 ounces and a standard deviation of 0.6.

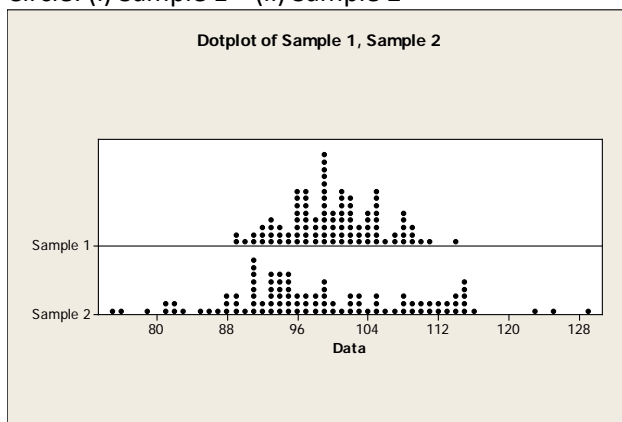
(13a. 3pts) Complete the null and alternative hypotheses by filling in the mathematical symbols.

$$H_0: \mu \quad 12$$

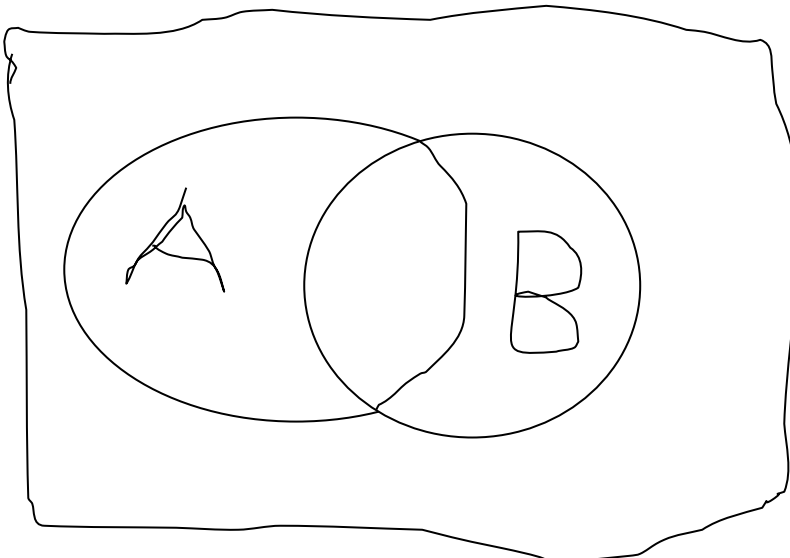
$$H_A: \mu \quad 12$$

(13b. 2pts) Calculate the t-statistic. Circle your answer.

(14. 3pts) Inspect the below dotplots. Each sample size is 100. Which sample has the **larger** variance?
Circle: (i) Sample 1 (ii) Sample 2



(15. 2pts) In the Venn diagram below, shade in the appropriate region for $A^c \cup B$.



(16.) Suppose the number of people in a 4 passenger car is distributed according to the following fictitious probability distribution. A car can contain as few as 1 person and as many as 4 people. The probability distribution for the number of people (X) is given below with the exception of the probability for 2 people.

K	P(X=x)	cdf: P(X≤x)
1	0.6	???=
2	????	???=
3	0.1	???=
4	0.1	???=

(16a. 3pts) Fill in the table for P(X=2). Show your work.

(16b. 2pts) Fill in the table for the cumulative distribution function (cdf) column.

(16c. 2pts) Calculate the expected value for X. Show your work.

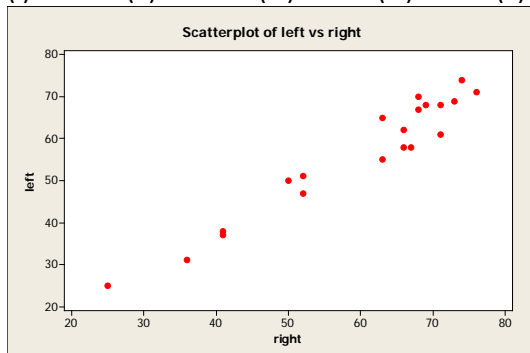
E(X) = _____

(17. 3pts) Suppose when determining a course grade, homework is weighted 40% and the exam 60%. If a person earned a 90% on homework and 70% on the exam, determine the final course grade in percent. Show your work.

Course grade= _____ %.

(18. 3pts) Circle the correlation that best fits the below graph?

(i) -2.00 (ii) -1.00 (iii) -0.97 (iv) -0.20 (v) 0 (vi) +0.20 (vii) +0.97 (viii) +1.00 (ix) +2.00



(19. 3pts) Calculate the probability of tossing three independent fair coins and getting three heads. Show your work.

Answer=_____. (Hint: This is easy.)

(20. 3pts) The probability of tossing 6 fair coins and getting 3 heads is **(Show work)**:

Circle one of the following: (i) 0.015625 (ii) 0.125 (iii) 0.3125 (iv) 0.5 (v) 0.6875 (vi) 0.75 (v) 0.5

(21. 4pts) You randomly survey 400 graduating seniors across the Cal State University system and ask them if are going to graduate school next year. 50 of them answer “yes”. Calculate a 95% confidence interval for the proportion of graduating seniors going to graduate school next year. Show your work.

Answer= (_____ , _____)

(22.) Suppose eye and hair color were recorded for 300 randomly selected people. Use the below contingency table for the eye and hair color data to estimate the requested probabilities.

	Green eyes	Blue eyes	Brown eyes	Total
Dark hair	10	20	80	110
Light hair	40	90	60	190
Total	50	110	140	300

(22a. 1pt) $P(\text{Blue eyes and Light hair}) =$ _____ (as an unreduced fraction)

(22b. 1pt) $P(\text{Dark hair} \mid \text{Brown eyes}) =$ _____ (as an unreduced fraction)

(22c. 1pt) $P(\text{Blue eyes}) =$ _____ (as an unreduced fraction)

(22d. 1pt) If hair color and eye color were independent, calculate the expected number of students with blue eyes and light hair. Show your work.

Expected number of blue-eyed light-haired students = _____

(23. 2pts) Suppose a histogram of a large dataset looks bell-shaped. According to the empirical rule, about what percent of the data should be within plus or minus one standard deviation from the mean?
_____ %

(24. 2pts) If random variable Z is normally distributed with mean 0 and standard deviation 1, calculate $P(-2 < Z < 2)$. Show your work.

Answer= _____

(25.) Suppose random variable X is normally distributed with mean 100 and standard deviation 10.

(25a. 2pts) Calculate $P(X > 90)$. Show your work.

Answer= _____

(25b. 2pts) Calculate $P(\bar{X} > 90)$ when the sample size is 4. Show your work.

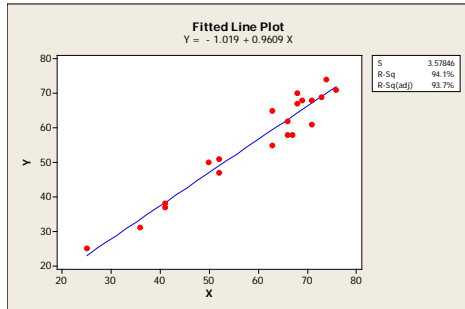
Answer= _____

(26) In the fall semester last year, there were 228 tenured or tenured track professors at Humboldt State of which 89 were female.

(26a. 2pt) If you were to draw a pie chart for these data, how many degrees would the slice for females be? (Show work)

(26b. 1pt) Sketch the pie chart.

(27.) Use the below least-squares regression equation $Y = -1.019 + 0.9609 X$ to answer the following questions.



(27a. 2pts) Calculate the predicted value of Y , \hat{Y} , if $X=30$. Show your work.

Answer= _____

(27b. 2pts) For each unit increase in X , we should expect the mean value of Y to

increase by _____ units.

(27c. 1pt) Circle which sentence best describes how the line is determined in least-squares regression.

- (i) The slope and intercept are chosen so that the sum of the closest distance from each point to the line is minimized.
- (ii) The slope and intercept are chosen so that the sum of the squared residuals is minimized.
- (iii) The slope and intercept are chosen so that the line touches the most points possible.

(28. 1pt) In the below graph there is one point at about (1.15, 64). In the below graph, draw a line representing the residual value for that point.

