Midterm #2

answers

Instructions: Each sub-question is worth 5 points. The exam is out of 80 points, so you get 20 points for free. Refer to the provided probability tables as necessary. You may use a calculator, and one sheet of notes. You will never be penalized for showing work, but if what is asked for can be computed directly, points awarded will depend primarily on the correctness of your numerical answer. Good luck!

Problem 1 The Capital Brewery in Middleton, WI produces a critically acclaimed seasonal Blonde Dopple-back brew. The beer is first available at Bockfest, an annual outdoor party which takes place in mid-February, in which 32-ounce collectable mugs of Blond Dobblebock are served.

Kirby wishes to estimate μ, the average number of Blonde Dopplebocks consumed by Bockfest attendees. To do so, he randomly samples 49 attendees at conclusion of Bockfest, and obtains the following information:

sample mean	$\overline{x}$	2.34
sample standard deviation	s	.67

a. Give a 97% confidence interval for μ.

$[2.13, 2.55]$

b. Bob Alice wants a confidence interval for μ which is no more than .02 wide. What is the maximum confidence level that Bob can use for such an interval?

This would be an 8.32% confidence interval, $[2.33, 2.35]$.

c. Suppose the true mean and standard deviation are $\mu = 2.2$ and $\sigma = .38$. What percentage of attendees will drink at least 3 mugs of Dobblebock?[1]

If amount of beer consumer is normally distributed, then about 1.76% of attendees will have three or more mugs of beer.

Problem 2 Hook's Cheese Company, in Mineral Point, WI, produces a 12-year aged cheddar which sells for about \$50/pound. Hook's has found that to age such a delicious cheese, the ideal ambient humidity is 70%, at a temperature of 54 degrees.

As part of Hook's quality control, the humidity in their storage cellar is randomly sampled several times per day. When reviewing the past week's data, Mr. Hook notices that the average humidity level in a sample of 36 observations is only 69%. He is concerned about the humidity falling under the target level. The sample deviation is .7.

a. Describe appropriate null and alternative hypotheses to test whether or not the cellar has appropriate humidity.

$$H_0 : \mu = 70\%$$
$$H_A : \mu < 70\%$$

[1] Do not attempt to drink 3 mugs of Blonde Dobblebock. It is 7.8% alcohol.

b. Describe the critical value of the test (the minimum t-stat for which you would reject H_0). Assume that $\alpha = .1$. Do you reject H_0?

For $\alpha = .1$, we reject H_0 if $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} < -1.28$. With the data collected, the test statistic is -8.57, so we reject H_0.

c. Determine the p-value of the test, and interpret its value in words.

The p-value is the lowest value of α for which H_0 would be rejected. Since the t-stat is -8.57, the p-value is a number very close to zero.

Problem 3 State Street is a mile-long pedestrian walkway in downtown Madison, WI. It stretches from the University of Wisconsin on the west to the state capital on the east, and is home to a variety of restaurants, bars, and shops.

Mayor Paul is reviewing police data on State Street deployments. Between 5 and 10 police officers patrol State Street each day, but the number varies from day to day. For each day in the previous 5 years, he has data on the number of police officers patrolling State Street, and the number of crimes that take place. He runs the following regression:

$$CRIME = \beta_0 + \beta_1 * POLICE + \epsilon \quad (1)$$

a. Explain in plain English how to interpret the coefficient β_1.

Given one more police officer deployed to State Street, β_1 is the change in expected number of crimes committed.

Suppose Mayor Paul obtains the following results:

Regression	**Statistics**
R Square	.204
Standard error	.98
Observations	1,825

	coefficients	Standard error	t stat	P-value	Lower 95%	Upper 95%
Intercept	8.40	3.53	2.38	.017	1.48	15.32
POLICE	1.64	.36	4.56	5.22E-6	.93	2.35

c. Is the POLICE variable significant, given a 5% chance of Type I error?

Yes, given such a low p-value, the police variable is significant even given a value of α of .01.

d. Mayor Paul is puzzled over the positive coefficient estimate β_1. Should he conclude that more police cause there to be more crime, and if not, why not?

This may be an issue of reverse causality; while police presumably affect the number of crimes committed, the number of crimes committed may also affect police staffing decisions. If it has been the case that more police are deployed to state street when crime there increases, then the regression results capture both effects, and so do not necessarily identify the causal effect of police on crime.

Problem 4 The Memorial Union in Madison, WI offers an idyllic setting for beer, food, and music on the shore of Lake Mendota. Memorial Union staff are interested in more accurately predicting how much beer they will sell as a function of the outdoor temperature. They employ the following regression model:

$$BEER = \beta_0 + \beta_1 * TEMPERATURE + \epsilon \tag{2}$$

where $BEER$ is daily kegs of beer sold, and $TEMPERATURE$ is the daily high temperature in degrees Farenheit.

The Memorial Union is open every day May-September when the daily high temperature is above 40 degrees. Staff have data from the previous 10 years.

Regression	**Statistics**
R Square	.312
Standard error	5.2198
Observations	1,476

	coefficients	Standard error	t stat	P-value	Lower 95%	Upper 95%
Intercept	-58.6	8.40	-6.98	3.03E-12	-75.06	-42.14
TEMPERATURE	1.57	.76	2.066	.039	.08	3.06

a. Interpret the estimates for β_0 and β_1 in words.

β_0 is the intercept; it has no economic significance, but is important so that the regression line fit as closely as possible to the data. β_1 is the marginal change in sales given a one-unit increase in temperature.

b. Consider the following hypothesis test:

$$H_0 : \beta_1 = 0$$
$$H_A : \beta_1 \neq 0$$

Is there sufficient evidence to support rejecting H_0, with $\alpha = .1$? What about $\alpha = .05$? What about $\alpha = .01$? How do you know?

Given the stated p-value of .039, $TEMPERATURE$ is significant at the $\alpha = .05$ level, but not the $\alpha = .01$ level.

c. Give a point estimate and an interval estimate for the amount of beer sold on a 78 degree day. Suppose that average temperature during the period of the dataset is 67 degrees, and that the variance of temperature is 19.2.

Point estimate: $Beer = -58.6+1.57*78 = 63.86$. Interval estimate: $63.86 \pm 1.96*5.22*\sqrt{\frac{1}{1476} + \frac{(78-67)^2}{1475*19.2}} = [63.25, 64.47]$. Since no confidence level was given, I used a 95% confidence interval. Other confidence levels are fine too.

Problem 5 Like the University of Kentucky, the University of Wisconsin is a land-grant institution, meaning that it received federal land under the Morrill Act of 1862 in support of a mission teaching agriculture, science, military science, and engineering. To commemorate this aspect of its history, a statue of President Lincoln sits at the top of Bascom Hill, at the center of campus.

Bo is interested in how many people walk in front of the statue. To this end, he randomly chooses 36 15-minute intervals during class hours, and observes the following:

sample mean	$\overline{x}$	282
sample standard deviation	s	51

a. Develop a 90% confidence interval for the number of people walking in front of the statue in each 15-minute time interval.

$[268.02, 295.98]$.

b. If the standard deviation remains the same, calculate the sample size necessary to obtain a 90% confidence interval that is no more than 8 in width.

440

c. Suppose Bo now wishes to estimate the number of people walking in front of the statue in each 1-hour interval. He considers two ways of obtaining this estimate:

1. Multiply the upper and lower limits of the confidence interval obtained in a. by 4.
2. Discard the above data and conduct a new survey, randomly sampling 36 1-hour periods.

Which of these two methods do you think will produce the *smaller* confidence interval, and why?

Method 2 will produce a smaller confidence interval. The simplest way of thinking about this is that method 2 will produce more data, and so must therefore produce tighter estimates.

Extra credit Predict the score of Saturday's Wisconsin-Kentucky Final Four game. The guess with the lowest mean squared deviation from the actual score will earn 3 extra points on the exam. For reference, I have provided scores in each team's previous NCAA tournament games:

Wisconsin	Kentucky
American University, 75-35	Kansas State, 56-49
Oregon, 85-77	Wichita State, 78-76
Baylor, 69-52	Louisville, 74-69
Arizona, 64-63	Michigan, 75-72

I will enter as well. If I win, I will deduct 3 points from each student's score. My prediction is that Wisconsin will defeat Kentucky, 203-6.

Your prediction: