Stat 2470, Exam #3, Spring 2014 Name KEY

Instructions: You may use your calculator for any functions the TI-84/83 model calculator is capable of using, such as probability distributions and obtaining graphs of any data. To show work on these problems, report the functions and their syntax as entered. Other things, such as integrating, must be done by hand unless specifically directed otherwise. Round means to one more place than the original data, variances and standard deviations to two more than the original data. Round probabilities to three significant figures or use exact values. In order to receive partial credit on any problem, you must show some work or I will have nothing to award partial credit on. Be sure to complete all the requested parts of each problem.

1. Tensile strength tests were carried out on two different grades of wire rod resulting in the accompanying data:

Grade	Sample Size	Sample Mean (kg/mm^2)	Sample St. Dev.
AISI 1064	$m = 130$	$\bar{x} = 108$	$s_1 = 1.3$
AISI 1078	$n = 130$	$\bar{y} = 124$	$s_2 = 2.0$

a. Does the data provide compelling evidence for concluding that true average strength for the 1078 grade exceeds that for the 1064 grade by more than 10 kg/mm^2? Test the appropriate hypotheses using the P-value approach. (9 points)

$H_0: \mu_1 - \mu_2 = -10$ $H_a: \mu_1 - \mu_2 < -10$

2-SampTTest

Set $\bar{x}_2 = 124 - 10 = 114$

& $\mu_1 < \mu_2 - 10$

$P \leq .001$

(calc reads 0)

reject the Ho.

the difference is at least 10 units

b. Estimate the difference between true average strengths for the two grades in a way that provides information about precision and reliability. [Hint: find a confidence interval.] (7 points)

CI for difference

$\bar{x} - \bar{y} = -16 \pm 2.61(.2092) \Rightarrow (-16.546, -15.454)$

$t_{\alpha/2, n-1} = 2.61$

$\sqrt{\frac{S_1^2}{130} + \frac{S_2^2}{130}} = .2092$

99% confidence interval

(your interval confidence level may vary)

or use 2-SampTInt

(also okay to use Z since sample size very large)

2. Suppose μ_1 and μ_2 are true mean stopping distances at 50 mph for cars of a certain type equipped with two different types of braking systems. Use the two-sample t test at significance level .01 to test $H_0: \mu_1 - \mu_2 = -10$ versus $H_a: \mu_1 - \mu_2 < -10$ for the following statistics: $m = 6$, $\bar{x} = 116$, $s_1 = 5.0$, $n = 6$, $\bar{y} = 129$, and $s_2 = 5.5$. (7 points)

2-SampTTest

$\bar{X}_2 = 129 - 10 = 119$

$P = .17318...$

$t = -.988...$

$P > .01$ $\therefore$ fail to reject Ho

There is not enough evidence to say the stopping distances are different by more than 10 feet

3. Two types of fish attractors, one made from vitrified clay pipes and the other from cement blocks and brush, were used during 16 different time periods spanning 4 years at Lake Tohopekaliga, Florida The following observations are of fish caught per fishing day.

Period

	1	2	3	4	5	6	7	8
Pipe	.00	1.80	4.86	.58	.37	.32	.11	.23
Brush	.48	2.33	5.38	.79	.32	.76	.52	.91

Period

	9	10	11	12	13	14	15	16
Pipe	.29	.85	6.64	.57	1.83	7.89	.63	.42
Brush	.75	1.61	9.73	.83	2.17	8.21	.56	.75

Does one attractor appear to be more effective on average than the other?

a. Use the paired *t* test with $\alpha = .01$ to test $H_o: \mu_D = 0$ versus $H_a: \mu_D \neq 0$. (8 points)

T-Test on difference $\bar{d} = -.544$ $\mu_{pipe} - \mu_{brush} = \bar{d}$

$S_D = .714$

$t = -3.049$ $p = .008 < .01$

reject Ho : brush is more effective than pipe

b. What happens if the two-sample *t* test is used ($s_1 = 2.48$ and $s_2 = 2.91$)? (7 points)

2 Samp T test (from data)

$t = -.569$

$p = .573$ fail to reject Ho

this method does not account for changing conditions

4. Ionizing radiation is being given increasing attention as a method for preserving horticultural products. A study reports that 153 of 180 irradiated garlic bulbs were marketable (no external sprouting, rotting, or softening) 240 days after treatment, whereas only 117 of 180 untreated bulbs were marketable after this length of time. Does this data suggest that ionizing radiation is beneficial as far as marketability is concerned? (6 points)

$\hat{p}_1 = \frac{153}{180}$ treated

$\hat{p}_2 = \frac{117}{180}$ untreated

Ho: $p_1 = p_2$

Ha: $p_1 \neq p_2$

2 PropZ test

$z = 4.38$

$p = 1.18 \times 10^{-5}$ < any reasonable α

reject Ho

5. Folacin is the only B vitamin present in any significant amount in tea, and recent advances in assay methods have made accurate determination of folacin content feasible. Consider the accompanying data on folacin content for randomly selected specimens of the four leading brands of green tea.

Brand	Observations					
1	8.0	6.3	6.7	8.7	9.0	10.2	9.7
2	5.8	7.6	9.9	6.2	8.5		
3	6.9	7.6	5.1	7.5	5.4	6.2	
4	6.5	7.2	8.0	4.6	5.1	4.1	

Does this data suggest that true average folacin content is the same for all brands?

a. Carry out a single-factor ANOVA test using $\alpha = .05$ via the *P*-value method with your calculator. (7 points)

ANOVA$(L_1, L_2, L_3, L_4) \Rightarrow F = 3.194$. $P = .0470 < .05$

df=3 $SSTr = 21.00$ $MSTr = 7.00$ $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

df=19 $SSE = 41.65$ $MSE = 2.19$ H_a: at least one μ_i different

reject H_0. one or more is different

b. Perform a multiple comparisons analysis to identify significant differences among brands using Tukey's method. (13 points)

from tables: $Q_{.05,4,20} = 3.96$

$w = 3.96\sqrt{\frac{2.09}{2}\left(\frac{1}{J_i}+\frac{1}{J_j}\right)}$

$J_1 = 7$, $J_2 = 5$

$J_3 = J_4 = 6$

get means from 1-VarStats (specify list)

$\bar{X}_1 = 8.32$

$\bar{X}_2 = 7.6$

$\bar{X}_3 = 6.45$

$\bar{X}_4 = 5.92$

4 3 2 1

$\bar{x}_i - \bar{x}_j$

↓

pairs

1,2	.72	± 2.37
1,3	1.87	± 2.25
1,4	2.4	± 2.25 ← only pair whose interval does not contain zero.
2,3	1.15	± 2.45
2,4	1.68	± 2.45
3,4	.53	± 2.34

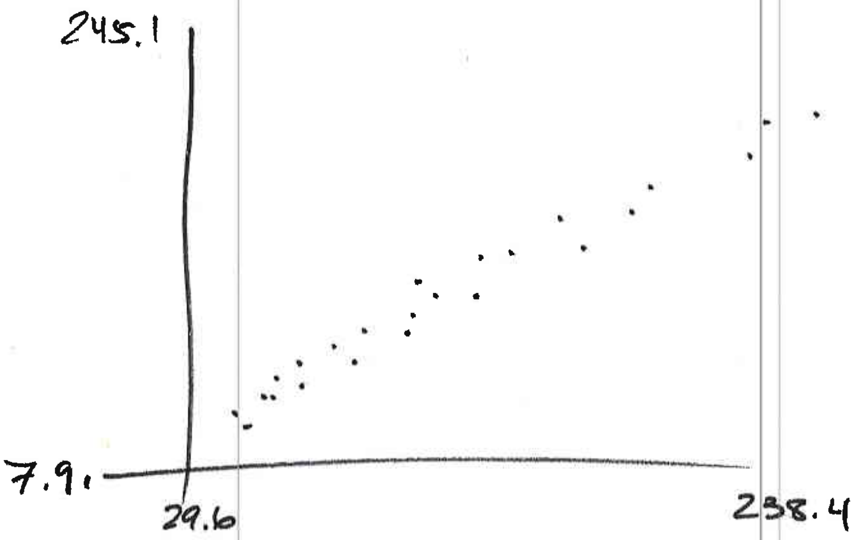
6. The accompanying observations on x = hydrogen concentration (ppm) using a gas chromatography method and y = concentration using a new sensor method were obtained in a recent study

x	47	62	65	70	70	78	95	100	114	118
y	38	62	53	67	84	79	93	106	117	116

x	124	127	140	140	140	150	152	164	198	221
y	127	114	134	139	142	156	149	154	200	215

a. Construct a scatter plot. Be sure to label the axes, and their range, clearly. (6 points)

Stat Plot
Zoom Stat



Something like this

b. Does there appear to be a very strong relationship between the two types of concentration measurements? Do the two methods appear to be measuring roughly the same quantity? Explain your reasoning. (6 points)

yes, there does appear to be a strong relationship the differences appear small & so may be only measurement errors

c. Construct the linear regression line for the data. (6 points)

$y = -.393 + .99067x$

d. How much of the data is explained by the relationship between the two variables? (4 points)

$r^2 = .9795$

almost 98%

7. A study contains a plot of the following data pairs, where x = pressure of extracted gas (microns) and y = extraction time (min):

x	40	130	155	160	260	275	325	370	420	480
y	2.5	3.0	3.1	3.3	3.7	4.1	4.3	4.8	5.0	5.4

a. Estimate σ and the standard deviation of $\hat{\beta}_1$. (6 points)

$S = .110$ $\quad S_{\hat{\beta}_1} = \dfrac{.110}{\sqrt{176,852}} = .000262$

b. Suppose the investigators had believed prior to the experiment that on average there would be an increase of .006 min. in extraction time associated with an increase of 1 micron in pressure. Use the P-value approach with a significance level of .10 to decide whether the data contradicts this prior belief. (8 points)

$H_0: \beta_1 = .006 \quad H_a: \beta_1 \neq .0060 \quad \alpha = .1$

$t = \dfrac{.0068 - .0060}{.000262} = 3.06$

tcdf$(3.06, E99, 8) = .00779 \times 2 = .01558 < .1$

for 2-tailed

reject H_0

8. Hydrogen content is conjectured to be an important factor in porosity of aluminum alloy castings. The accompanying data on x = content and y = gas porosity for one particular measurement technique have been reported:

x	.18	.20	.21	.21	.21	.22	.23
y	.46	.70	.41	.45	.55	.44	.24

x	.23	.24	.24	.25	.28	.30	.37
y	.47	.22	.80	.88	.70	.72	.75

MINITAB gives the following output in response to a CORRELATION command:

Correlation of Hydrogen and Porosity = 0.449

a. Test at level .05 to see whether the population correlation coefficient differs from 0. (8 points)

LinRegT Test easiest for this

$t = 1.74 \quad p = .107$

$H_0: \rho = 0, \quad H_a: \rho \neq 0$ there is not sufficient evidence to reject H_0

b. If a simple linear regression analysis had been carried out, what percentage of observed variation in porosity could be attributed to the model relationship? (4 points)

About 20% $\quad r^2 = .2017$

9. An investigation of the influence of sodium benzoate concentration on the critical minimum pH necessary for the inhibition of Fe yielded the accompanying data, which suggests that expected critical minimum pH is linearly related to the natural logarithm of concentrate:

Concentration	.01	.025	.1	.95	X
pH	5.1	5.5	6.1	7.3	Y

a. What is the implied probabilistic model, and what are the estimates of the model parameters? [Hint: create a linear model with ln(x).] (8 points)

$y = 7.2867 + .483 \ln(x)$

b. What critical minimum pH would you predict for a concentration of 1.0? Obtain a 95% PI for critical minimum pH when concentration is 1.0. (6 points)

$\hat{y} = 7.2867$ $\quad \ln(1) = 0$

$t_{\alpha/2, 2} = 4.303$

$s = .06496$

from LinRegTTest

$\overline{(\ln x)} = -2.66$

$7.29 \pm 4.303 * .06496 \sqrt{1 + \frac{1}{4} + \frac{(-2.66)^2}{40.120 - \frac{(10.648)^2}{4}}}$

$= 7.29 \pm 4.303 * .06496 * 1.360$

$= 7.29 \pm .38$

$(6.91, 7.67)$ is the PI

10. A trucking company considered a multiple regression model for relating the dependent variable y=total daily travel time for one of its drivers (hours) to the predictors x_1 =distance traveled (miles) and x_2 the number of deliveries made. Suppose that the model equation is

$$Y = -.800 + .062x_1 + .900x_2 + \varepsilon$$

a. What is the mean value of travel time when distance traveled is 50 miles and three deliveries are made? (6 points)

$\hat{y} = -.800 + .062(50) + .90(3) = 5$ hours

b. How would interpret $\beta_1 = .060$, the coefficient of the predictor x_1? What is the interpretation of $\beta_2 = .900$? (6 points)

for each addition mile added to distance traveled, the time increased by approximately .06 hours. or about 3.6 minutes

like for each delivery, time increases by approximately .9 hours or about 54 minutes

c. If $\sigma = .5$ hour, what is the probability that travel time will be at most 6 hours when three deliveries are made and the distance traveled is 50 miles? (5 points)

$\bar{y} = 4.9, \quad \sigma = .5$

$z = \dfrac{6 - 4.9}{.5} = 2.2$

$P(z \leq 2.20)$ = normalcdf$(-E99, 2.20) = .9861$

about 98.6% of travel times will be less than 6 hours