Stat 2470, Final Exam, 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. What conclusion would be appropriate for an upper-tailed chi-squared test in each of the following situations? (5 points each)

   a. $\alpha = .05, df = 4, \chi^2 = 14.69$ $.005389 < .05$ reject Ho

   b. $\alpha = .01, df = 3, \chi^2 = 7.85$ $.04921 > .01$ fail to reject Ho

   c. $\alpha = .10, df = 2, \chi^2 = 4.42$ $.1097 > .10$ fail to reject Ho

   d. $\alpha = .01, k = 6, \chi^2 = 13.92$ $.03054 > .01$ fail to reject Ho

2. A statistics department at a state university maintains a tutoring service for students in its introductory service courses. The service has been staffed with the expectation that 40% of its students would be from the business statistics course, 30% from engineering statistics, 20% from the statistics course for social science students, and the other 10% from the course for agriculture students. A random sample of n=120 students revealed 50, 40, 18, and 12 from the four courses. Does this data suggest that the percentages on which staffing was based are not correct? State and test the relevant hypotheses using $\alpha = .05$. (9 points)

	bus stat(1)	eng stat(2)	soc sci(3)	other(4)	
observed	50	40	18	12	
expected	48	36	24	12	all at least 5

$\chi^2 \quad \frac{2^2}{48} + \frac{4^2}{36} + \frac{6^2}{24} + \frac{0^2}{12} = \frac{4}{48} + \frac{16}{36} + \frac{36}{24} + 0 = 2.0278 \quad df = 3$

χ^2 p-value $= .5666 > .05$ The staffing decisions were fine

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 $H_a : \mu_D \neq 0$. (10 points)

T-Test on differences

$\mu_{pipe} - \mu_{brush} = \bar{d}$

$\bar{d} = -.544$

$p = .008 < .01$

reject Ho

$S_D = .714$

brush is more effective than pipe

$t = -3.049$

b. What happens if the two-sample t test is used ($s_1 = 2.48$ and $s_2 = 2.91$)? (8 points)

2-SampTTest

this method does not account for changing conditions

$t = -.569$

$p = .573$

fail to reject Ho

4. 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. (10 points)

LinRegTTest easiest for this

$t = 1.74$ $p = .107$

$H_0: \rho = 0$ $H_a: \rho \neq 0$

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? (6 points)

about 20% $r^2 = .2017$

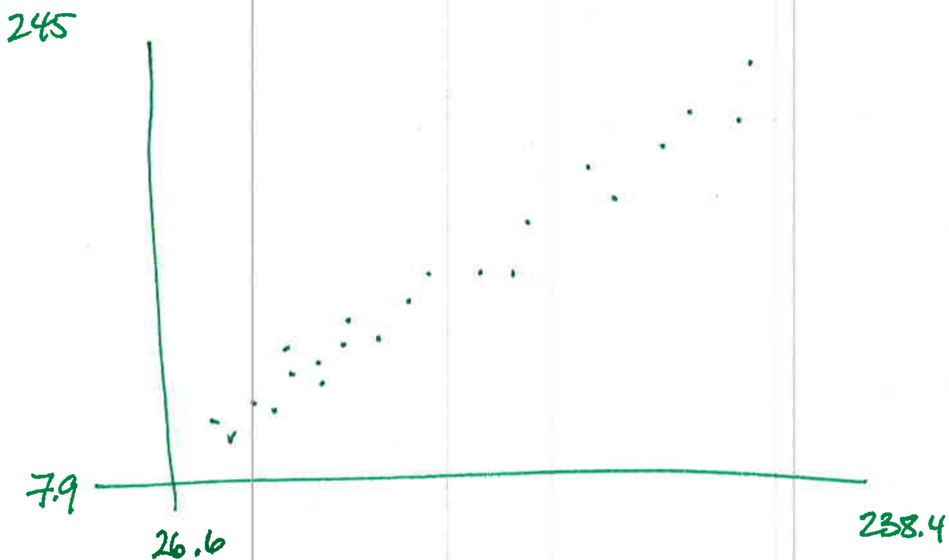
5. The accompanying observations on x = hydrogen concentration (ppm) using a gas chromatography method and y = concentration using a new sensor method were obtained in a recent study

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. (10 points)

(use the space on the next page)



from calculator
StatPlot
Zoom Stat

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. (7 points)

Yes, there does appear to be a strong relationship. the differences appear small and so may be only measurement errors

c. Construct the linear regression line for the data. (8 points)

$y = -.393 + .99067x$

d. How much of the data is explained by the relationship between the two variables? (5 points)

$r^2 = .9795$

almost 98%

6. A certain pair of events if governed by the joint probability distribution $f(x,y) = \begin{cases} \frac{6}{7}(xy+y^2), 0 \le x \le 1, 0 \le y \le 2x \\ 0, \quad\quad\quad otherwise \end{cases}$

a. Verify that this is a valid probability distribution. (8 points)

$$\frac{6}{7}\int_0^1\int_0^{2x} xy + y^2\,dy\,dx = \frac{6}{7}\int_0^1 \frac{xy^2}{2} + \frac{y^3}{3}\Big|_0^{2x} dx =$$

$$\frac{6}{7}\int_0^1 \frac{x(2x)^2}{2} + \frac{(2x)^3}{3}dx = \frac{6}{7}\int_0^1 2x^3 + \frac{8}{3}x^3\,dx = \frac{6}{7}\int_0^1 \frac{14}{3}x^3\,dx$$

$$= \frac{6}{7}\left[\frac{14}{3}\cdot\frac{x^4}{4}\Big|_0^1 = 1\right.$$ yes, it is a valid distribution

b. Find $Cov(X,Y)$. (15 points) $= E(XY) - E(X)E(Y)$

$$E(XY) = \frac{6}{7}\int_0^1\int_0^{2x}(xy+y^2)xy\,dy\,dx = \frac{6}{7}\int_0^1\int_0^{2x} x^2y^2 + xy^3\,dy\,dx = \frac{6}{7}\int_0^1 \frac{x^2y^3}{3} + \frac{xy^4}{4}\Big|_0^{2x}dx$$

$$= \frac{6}{7}\int_0^1 \frac{20}{3}x^5\,dx = \frac{6}{7}\cdot\frac{20}{3}\cdot\frac{1}{6}x^6\Big|_0^1 = \frac{20}{21}$$

$$E(X) = \frac{6}{7}\int_0^1\int_0^{2x} x^2y + xy^2\,dy\,dx = \frac{6}{7}\int_0^1 \frac{x^2y^2}{2} + \frac{xy^3}{3}\Big|_0^{2x} = \frac{6}{7}\int_0 \frac{14}{3}x^4\,dx =$$

$$\frac{6}{7}\cdot\frac{14}{3}\cdot\frac{1}{5}x^5\Big|_0^1 = \frac{4}{5}$$

$$E(Y) = \frac{6}{7}\int_0^1\int_0^{2x} xy^2 + y^3\,dy\,dx =$$

$$\frac{6}{7}\int_0^1 \frac{xy^3}{3} + \frac{y^4}{4}\Big|_0^{2x}dx = \frac{6}{7}\int_0^1 \frac{20}{3}x^4\,dx = \frac{6}{7}\cdot\frac{20}{3}\cdot\frac{1}{5}x^5\Big|_0^1 = \frac{8}{7}$$

$$Cov(X,Y) = \frac{20}{21} - \frac{4}{5}\cdot\frac{8}{7} = \frac{20}{21} - \frac{32}{35} = \frac{4}{105}$$

c. Is the distribution independent? Why or why not? (4 points)

no. the $Cov(X,Y) \neq 0$

$\$$ $f(x,y) \neq f_X(x) \cdot f_Y(y)$

7. Suppose that we have collected 7 pieces of data from an exponential distribution and obtained the following sample results: $7, 11, 19, 22, 24, 30, 38$. We'd like to use this data to estimate the value of the parameter λ.

 a. What is the maximum likelihood function for this data? (7 points)

$$f(\lambda) = \lambda e^{-7\lambda} \lambda e^{-11\lambda} \lambda e^{-19\lambda} \lambda e^{-22\lambda} \lambda e^{-24\lambda} \lambda e^{-30\lambda} \lambda e^{-38\lambda}$$

$$= \lambda^7 e^{-151\lambda}$$

 b. Use this function to estimate $\hat{\lambda}$. (You must show the calculus on this problem! You will receive no more than 1 point for correctly calculating the mean by another method.) (10 points)

$$f'(\lambda) = 7\lambda^6 e^{-151\lambda} - 151\lambda^7 e^{-151\lambda} = \lambda^6 e^{-151\lambda}(7 - 151\lambda) = 0$$

$$7 - 151\lambda = 0 \Rightarrow 7 = 151\lambda \Rightarrow \lambda = \frac{7}{151} \approx .046$$

$$\mu = \frac{1}{\lambda} \text{ so } \hat{\mu} = \frac{1}{\hat{\lambda}} = \frac{151}{7} \approx 21.57$$

8. It was reported that, in a sample of 1507 adult Americans, only 432 correctly described the Bill of Rights as the first ten amendments to the U.S. Constitution. Calculate a (two-sided) confidence interval using a 95% confidence level for the proportion of all U. S. adults that could give a correct description of the Bill of Rights. (8 points)

$N = 1507$

$\hat{p} = \frac{432}{1507} \approx .2866$

1 Prop Z Interval

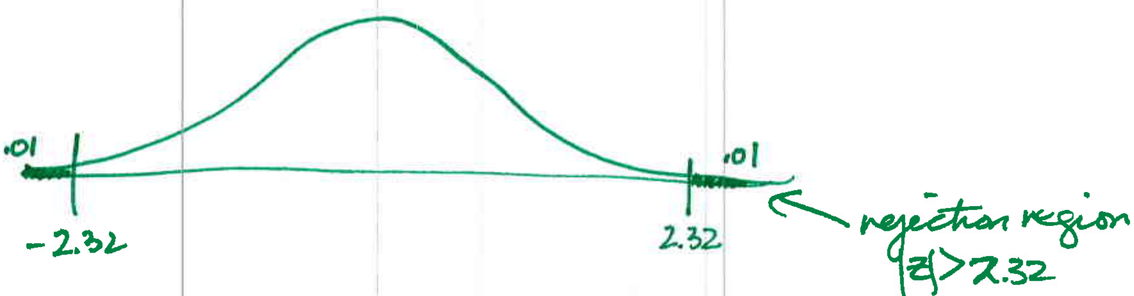
$(.26383, .30949)$

9. Explain what is the difference between a Type I and a Type II error. Give an example of each. (8 points)

a Type I error occurs when we reject the null hypothesis even though it is true: in a criminal case if someone is not lying, but we believe they are (not lying is the default)

a Type II error occurs when we fail to reject the null hypothesis even when it is false: in a lie detector test, the person is lying, but the test does not correctly pick it up.

10. Suppose that you want to conduct a hypothesis test on a mean from a normally distributed population with a unknown standard deviation, using a sample size of n=22. Sketch the normal curve and the rejection region for a two-tailed hypothesis test with $\alpha = 0.02$. Be sure to label the critical values and the rejection region clearly. (8 points)



11. Let μ denote the mean reaction time to a certain stimulus. For a large-sample *z* test of $H_0: \mu = 8$ versus $H_a: \mu > 8$, find the *P*-value associated with each of the given values of the *z* test statistics. (4 points each)

a. 1.52

.0643

b. 0.95

.1711

c. 0.79

.2148

12. A sample of 20 glass bottles of a particular type was selected, and the internal pressure strength of each bottle was determined. Consider the following partial sample information:

Median = 202.2
lower fourth = 196.0 Q_1
Upper fourth = 216.8 Q_3

extreme outlier smaller than 133.6

Three smallest observations	$\boxed{125.8}$	188.1	193.7
Three largest observations	221.3	230.5	$\boxed{250.2}$

above 248, below 279.2
mild outlier

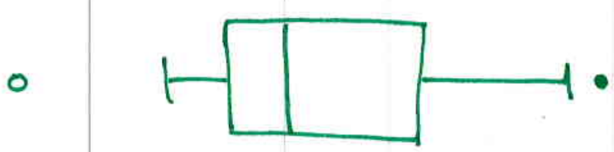
a. Are there any outliers in the sample? (7 points)

$1.5\ IQR = 31.2$ $IQR = 20.8$
$3\ IQR = 62.4$

mild outliers $196 - 31.2 = 164.6$
$216.8 + 31.2 = 248$

extreme outliers
$196 - 62.4 = 133.6$
$216.8 + 62.4 = 279.2$

b. Sketch a boxplot of the data. (8 points)



13. Let A denote the event that the next item checked out at a college library is a math book, and let B be the event that the next item checked out is a history book. Suppose that P(A) = .03 and P(B) = .05. The probability that the book is a history of math book is $P(A \cap B) = 0.001$.
   a. Why is it not the case that P(A) + P(B) = 1? (5 points)

   there are other books than math & history

   b. Calculate P(A') (4 points)

   $1 - .03 = .97$

   c. Calculate P($A \cup B$). (5 points)

   $.03 + .05 - .001 = .079$

   d. Calculate P($A' \cap B'$). (5 points)

   $1 - .079 = .921$

14. Suppose that an organization has 15 members, 7 of whom are male, and 8 of whom are female. Answer the following questions based on this information. (5 points each)
   a. How many possible combinations of President, Vice President and Secretary are there?

   $15P3 = 2730$

   b. How many possible combinations of President, Vice President and Secretary are all women?

   $8P3 = 336$

   c. If the officers are chosen at random, what is the probability that the officers will all be women?

   $\frac{336}{2730} \approx .123$ or 12.3%

d. If they wish to select a subcommittee to plan events for the coming year made up of four people, how many possible committees are there with exactly two men and two women?

$\binom{7}{2}\binom{8}{2} = 588$

15. Consider the following information: where A = {Visa Card}, B = {MasterCard}, P(A) = .5, P(B) = .4, and $P(A \cap B) = .25$. Calculate each of the following probabilities.

a. $P(B|A)$ (5 points)

$\frac{P(A \cap B)}{P(A)} = \frac{.25}{.5} = .5$

b. $P(A|B)$ (5 points)

$\frac{P(A \cap B)}{P(B)} = \frac{.25}{.4} = .625$

c. Given that an individual is selected at random and that he or she has at least one card, what is the probability that he or she has a Visa card? (5 points)

$P(A \cup B) = .5 + .4 - .25 = .65$

$P(A \mid A \cup B) = \frac{P[A \cap (A \cup B)]}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = \frac{.5}{.65} = .7692$

16. Suppose that only 16% of all drivers come to a complete stop at an intersection having flashing red lights in all directions when no other cars are visible. What is the probability that, of 20 randomly chosen drivers coming to an intersection under these conditions,

a. At most 4 will come to a complete stop? (5 points)

$p = .16 \quad n = 20$

binomialcdf(20, .16, 4) = .794

b. Exactly 4 will come to a complete stop? (5 points)

binomialpdf(20, .16, 4) = .195

c. At least 5 will come to a complete stop? (5 points)

1 - binomialcdf (20, .16, 4) = .2059

d. How many of the next 20 drivers do you expect to come to a complete stop? (6 points)

$E(X) = np$

$20 * .16 = 3.2$

17. Let X have a standard gamma distribution with $\alpha = 6$. (Assume β=1.) Evaluate the following: (4 points each)

a. $P(X \leq 5)$ .384

b. $P(X > 8)$ .191

c. $P(3 \leq X \leq 8)$ .725