STA 3024 **Exam 3** **Practice Problems**

NOTE: These are just Practice Problems. This is NOT meant to look just like the test, and it is NOT the only thing that you should study. Make sure you know all the material from the notes, quizzes, suggested homework and the corresponding chapters in the book.

Questions 1 – 7 Former kicker for the Gator football team, Chris Hetland, was very good at making field goals in the 2005 season, but in the 2006 regular season had only made 3 out of 12. The following is the Logistic Regression Output to predict the probability of making a field goal (yes/no), based on how far the kick is (in yards) and the year (2005 or 2006).

log. reg: predict prob. or prop. explained by cont. variable

```
Logistic Regression Table

Predictor       Coef     SE Coef       Z       P
Constant     8312.97     3073.50    2.70   0.007
yards      -0.173760   0.0901421   -1.93   0.054
year        -4.14174     1.53141   -2.70   0.007
```

· categorical (yes/no) response

1. What kind of variables do we have here?
a) a quantitative predictor and a quantitative response
b) two quantitative predictors and a quantitative response
c) a quantitative predictor and a categorical response
d) two quantitative predictors and a categorical response

$\hat{p} = \frac{e^{x}}{1+e^{x}}$ $\quad x = a + b_1x_1 + b_2x_2 + \varepsilon$

$\hat{p} = \frac{e^{8312.97 - 0.173760(\text{yds}) - 4.14174(\text{yr})}}{1 + e^{8312.97 - 0.173760(\text{yds}) - 4.14174(\text{yr})}}$

2. Write down the fitted logistic regression equation:
a) $\hat{y} = 8312.97 - 0.173760 - 4.14174$
b) $\hat{y} = 8312.97 - 0.173760 \text{ yards} - 4.14174 \text{ year}$
c) $\hat{p} = e^{8312.97 - 0.173760 - 4.14174} / 1 + e^{8312.97 - 0.173760 - 4.14174}$
d) $\hat{p} = e^{8312.97 - 0.173760 \text{ yards} - 4.14174 \text{ year}} / 1 + e^{8312.97 - 0.173760 \text{ yards} - 4.14174 \text{ year}}$

neg. exponent ↓ smaller value out of formula ↓ smaller $\hat{p}$

3. The coefficients of yards and years are both negative. This means:
a) that neither variable is a good predictor of whether the kicker will make or not the field goal
b) that simple linear regression would have been more appropriate than logistic regression
c) that there was a mistake in the way the data was entered into the computer
d) that the chances of making the field goal go down as the yardage increases, and as the years increase

Find the probability of making a field goal:

4. from the 30 yd line in 2006
5. from the 30 yd line in 2005
6. from the 40 yd line in 2006
7. from the 40 yd line in 2005

4) $\hat{p} = \frac{e^{8312.97 - 0.173760(30) - 4.14174(2006)}}{1 + e^{8312.97 - 0.173760(30) - 4.14174(2006)}}$

exp: $8312.97 - 0.173760(30) - 4.14174(2006) = -0.57324$

$e^{\text{exp}} : e^{-0.57324} = 0.56361961$

$\hat{p} = \frac{0.56}{1+0.56} = 0.36$

5) exp: $8312.97 - 0.173760(30) - 4.14174(2005) = 3.56685$

$e^{3.56685} = 35.46233582$

$\hat{p} = \frac{35.46}{1+35.46} = 0.97258$

6) exp: $8312.97 - 0.173760(40) - 4.14174(2006) = -2.31084$

$e^{-2.31084} = 0.09917790 71$

$\hat{p} = \frac{0.09917790 71}{1+0.0992} = 0.0902$

7) exp: $8312.97 - 0.173760(40) - 4.14174(2005) = 1.8309$

$e^{1.8309} = 6.23934$

$\hat{p} = \frac{6.23935}{1+6.23935} = 0.862$

Question 8 - 13 As part of a project for their Intro Stat course, two students compared two brands of chips, Frito Lays and Golden Flakes, to see which company gives you more for your money. Five bags of each brand (which, according to the label, each contained 35.4 grams) were measured with a very accurate scale. Use the Wilcoxon Rank-Sum test to see if there are any significant differences between the two brands in the amount of product they put in their bags. $=$ vs $\neq$

	1.5	3	4	5.5	5.5
Frito Lays:	35.3	35.4	35.8	35.9	35.9
	1.5	7	9	8	10
Golden Flake:	35.3	37.8	38.8	38.1	42.5

35.3 : ranks 1 & 2, avg = 1.5

35.9 : ranks 5 & 6, avg = 5.5

outlier - bigger than others

non-parametric → medians or distributions

8. The null hypothesis is about:
a) the mean contents of the bags for Frito Lays and Golden Flakes brands
b) the mode of the contents of the bags for Frito Lays and Golden Flakes brands
c) the distribution of the contents of the bags for the two brands
d) the number of bags with contents below the label weight for the two brands – ~~sign~~ test for median

9. The alternative hypothesis, according to the problem stated above, is that:
a) Frito Lays gives you more chips than Golden Flakes
b) Frito Lays gives you less chips than Golden Flakes
c) Frito Lays gives you either more or less chips than Golden Flakes
d) Golden Flakes gives you more chips than the amount stated on the label

10. The bags that contained 35.9 grams will receive a rank of:
a) 4 b) 4.5
c) 5 d) 5.5 e) 6

11. The p-value for the test was .1164. We conclude that:
a) Frito Lays gives you more chips.
b) Golden Flakes gives you more chips.
c) There is not enough evidence to prove a difference between the two brands.
d) There is enough evidence to prove a difference between the two brands.

compare to $\alpha = 0.1, 0.05, 0.01$
. p-val > all α
↓
fail to rej. H_0

12. If the assumptions for the Normal based procedure were satisfied, we could analyze the data with a confidence interval for:
a) μ ↑ single mean
b) $\mu_1-\mu_2$ ↑ 2 sample t-test (indep)
c) μ_d ↑ 2 dependent means
d) $\eta_1-\eta_2$ ↑ Wilcoxon Rank Sum test

13. Why is it not a good idea to use the Normal-based procedure here?
a) the data was not randomly selected
b) the data does not have a continuous distribution
c) the outlier violates the assumption of Normality
d) the nonparametric method is always better

Questions 14 - 18 Do plain and peanut m&m's have the same distribution of colors? Several bags of each variety (plain and peanut) were randomly selected, and the number of candies of each color were counted before eating any of them. The data appears below.

- expected values
- residuals

	brown	yellow	red	blue	green	orange	
plain	81 (74.26, +6.7)	84 (68.85, +15)	41 (51.53, −10.5)	17 (22.73, −5.7)	30 (33.34, −3.3)	41 (43.19, −7.2)	294
peanut	17 (23.74, −6.7)	7 (22.05, −15)	27 (16.47, +10.5)	13 (7.27, +5.7)	14 (10.66, +3.3)	16 (13.81, +7.2)	94
	98	91	68	30	44	57	388

more of 2 colors than expect

more of 41 colors than expect

∴ peanut more colorful than expected

14. The null hypothesis is that:
a) plain and peanut varieties are independent
b) the colors are independent of each other
c) color and variety are independent of each other
d) all of the above

two variables are independent
color type

15. The expected number of blue, peanut m&m's (under independence) is:
a) 15.0 b) 15.67
c) 7.27 d) 32.33

$\frac{\text{row total} \times \text{col. total}}{\text{grand total}}$

total blue = 30
total peanut = 94
grand total = 388

expect: $\frac{30 \times 94}{388} = 7.27$

16. The sampling distribution is χ^2 with degrees of freedom equal to:
a) 5 b) 10 c) 11 d) 12

(# rows − 1) × (# col. − 1)
(2 − 1) × (6 − 1)
1 × 5

17. The test statistic was 32.67. Use the table to approximate the p-value for this test:
a) smaller than .001
b) equal to .05
c) between .05 and .10
d) between .950 and .975

18. What conclusions can you reach from this analysis, based on the data and the test statistic given?
a) There is something wrong with the data, maybe the m&m's were not randomly selected.
b) The distribution of colors is not significantly different for plain and peanut m&m's.
c) Peanut m&m's are significantly more colorful than plain m&m's.
d) There are significantly more brown m&m's than orange m&m's.

Questions 19 - 21 Match each of the Nonparametric procedures presented on the left with the corresponding experimental design from the list on the right (use each alternative only once).

C 19. Kruskal-Wallis H Test (like anova)
A 20. Wilcoxon Rank-Sum Test (2 indep groups)
B 21. Wilcoxon Signed-Rank Test

a) two independent samples
b) paired samples
c) several independent samples

Questions 22 -24 Five sets of identical twins were selected at random from a population of identical twins. One child was selected at random from each pair to form an "experimental group." These five children were sent to school. The other five children were kept at home as a control group. At the end of the school year the following IQ scores were obtained. Does this evidence justify the conclusion that lack of school experience has a depressing effect on IQ scores? Analyze the data with the Wilcoxon Signed-Rank Test

(matched pairs)

Pair	Experimental Group	Control Group	Difference	abs. value diff	rank
1	110	112	-2	2	2 (-)
2	125	120	+5	5	3 (+)
3	139	128	+11	11	5 (+)
4	142	135	+7	7	4 (+)
5	127	126	+1	1	1 (+)

22. The sums of the ranks for this test are:

a) W+ =13 W- =2
b) W+ =24 W- =2
c) W =9
d) W =30

W- : 2
W+: 3+5+4+1 = 13

23. The data shows some evidence that:

a) the experimental (school) group tends to have higher IQs than the control (home)group.
b) the experimental (school) group tends to have lower IQs than the control (home)group.
c) the experimental (school) group tends to have IQs similar to the control (home)group.
d) the experimental (school) group tends to have IQs different from the control (home)group.

no neg. differences

24. Which of the following (one-sided) p-values looks reasonable for this data?

a) 0.0001
b) 0.9663
c) 0.0885
d) 0.4367

rej. @ at least one α-value
$\alpha = 0.1, 0.05, 0.01$
not super strong evidence

Questions 25 – 29 Data collected to study the relationship between child obesity and parental obesity is shown in the following contingency table.

- contribution to χ^2
- expected values

		Child: Obese	Child: Nonobese	total
Parent	Obese	34 (31.5) (0.1984)	29 (31.5) (0.1984)	63
	Nonobese	16 (18.5) (0.3378)	21 (18.5) (0.3378)	37
	total	50	50	100

25. What is the null hypothesis being tested? variables = child obesity & parent obesity

a) the proportion of obese and nonobese parents are the same
b) the proportion of obese and nonobese children are the same
c) the proportion of obese children is the same for obese and nonobese parents
d) all of the above

26. How many obese children were involved in the study?

a) 34 b) 50 c) 63 d) 100

27. What are the expected counts for each category under the null hypothesis (in the same order as the given table)? child obese: $\frac{50 \times 63}{100} = 31.5$ CN, PO: $\frac{50 \times 63}{100} = 31.5$ CO, PN: $\frac{50 \times 37}{100} = 18.5$ CN, PN: $\frac{50 \times 37}{100} = 18.5$

a) 34 29
16 21

b) 35 30
15 20

c) 31.5 31.5
18.5 18.5

d) 22.1 22.1
22.1 22.1

28. How many degrees of freedom are associated with the X^2 test?

a) 4 b) 3 c) 2 d) 1

row: 2
col: 2
$(2-1)\times(2-1)$
1

29. Find the contribution to the Test Statistc of the parent obese/child obese cell.

$\frac{(\text{obs} - \text{expect})^2}{\text{expected}}$ obs: 34 expect: 31.5 $\rightarrow \frac{(34 - 31.5)^2}{31.5}$

$= 0.1984$

arbetrary response → non-parametric

Questions 30 – 32 An experiment was conducted to determine whether a test designed to identify a certain form of mental illness could be easily interpreted with little psychological training. The test was given to 100 people (half of which had the illness, and half didn't) and fifteen people were asked to evaluate them. The fifteen judges were five staff members of a mental hospital, five trainees at the hospital, and five undergraduate psychology majors. The results in the table give the number of the 100 tests correctly classified by each judge. Analyze the data with the Kruskal-Wallis Test.

Staff		Trainees		Students	
78	8.5	80	12	65	1
76	7	69	2	74	4
80	12	75	5.5	78	8.5
79	10	81	14	80	12
86	15	72	3	75	5.5

75: ranks 5 & 6, avg = 5.5

77: ranks 8 & 9, avg = 8.5

80: ranks 11, 12, & 13, avg = 12

30. The ranks for the observations on the first row should be:
a) 2 3 1
b) 7 9 1
c) 8.5 12 1
d) none of the above

31. The highest rank given to any observation is:
a) 11 b) 5 c) 15 d) 3

32. If the p-value of the test is small we would conclude that there are:
a) differences between staff, trainees and students in their ability to interpret the test
b) no differences between staff, trainees and students in their ability to interpret the test
c) differences in the individual judges abilities to interpret the test
d) no differences in the individual judges abilities to interpret the test

Questions 33 – 35 **Q**uestions regarding the use of Nonparametric procedures:

33. Which of the following kinds of data <u>can</u> be analyzed with Nonparametric procedures?
a) normal ✓ (not best test) b) continuous ✓ c) ranks ✓ (what is normally used) d) all of the above

34. Which of the following kinds of data <u>should</u> be analyzed with Nonparametric procedures?
a) normal b) continuous c) ranks d) all of the above

35. Given that all the necessary assumptions for each test are satisfied, which are more powerful at finding significant differences?
a) Nonparametric procedures, since their assumptions are generally easier to satisfy.
b) Normal-based procedures, since they take into consideration the shape of the distribution.
c) Nonparametric procedures, since their assumptions are generally harder to satisfy.
d) Normal-based procedures, since they work for distributions of almost any shape.

normal: harder assumptions • consider shape of distribution

nonparametric: few assumptions • distribution of any shape

Questions 36 - 40 For each of the following stories, determine which would be the simplest type of statistical analysis that would be appropriate to use. Use each type of analysis only once.

a) Paired t test → 2 dep. means
b) Two sample t-test → 2 indep. means
c) ANOVA → 3+ indep. means
d) Kruskal-Wallis → 3+ indep. means & not normal
e) Wilcoxon Rank-Sum Test → 2+ indep. means & not normal

C 36. Compare the average number of hours per week spent on Facebook for Freshmen, Sophomore, Juniors and Seniors at UF, based on a random sample of 100 students. [4 groups] → anova / KW
[large sample] → normal → anova

D 37. Compare the distribution of the number of hours per week spent on Facebook for Freshmen, Sophomore, Juniors and Seniors at UF, based on random samples of 10 students per group, which had quite different standard deviations. [4 groups] → anova / KW
[small sample] → NOT NORMAL → KW

A 38. Compare the average number of hours per week spent on Facebook during the first week in April and the first week in May (finals week) for random students at UF, measured on the same 100 students. dep. groups → paired t-test or wil. SIGN rank test
big sample → normal
X don't have option

E 39. Compare the distribution of the number of hours per week spent on Facebook for male and female students at UF, based on a random sample of 10 students. There was an outlier in one of the groups. 2 indp. groups → t-test or wil RANK-sum
10 students → small sample → NOT NORMAL → rank-sum

B 40. Compare the average number of hours per week spent on Facebook for male and female students at UF, based on a random sample of 100 students. 2 indp groups → t-test or Rank sum
100 = sample-size → normal → t-test

Questions 41 - 45 For each of the following stories, determine which would be the simplest type of statistical analysis that would be appropriate to use. Use each type of analysis only once.

a) Confidence Interval for One Proportion → 1 group (1 cat. variable)
b) Contingency Table → 2 cat. variables predict 1 quant. variable
c) Simple Linear Regression → 1 quant var predict 1 quant variable
d) Multiple Regression → 2+ quant./cat. var predict 1 quant. var
e) Logistic Regression → 1+ quant/cat predict 1 cat. variable

D 41. Predict the average number of hours per week UF students spend on Facebook, based on their age and gender. predict: quant
explained by 2+ cat → ~~contingency table~~
mult. regression

A 42. Estimate the fraction of UF students who have Facebook accounts.
predict cat. of 1 group → one prop.

B 43. Determine if the fraction of UF students who have Facebook accounts is different for Males and Females. predict cat. var
compare % → cont. table

E 44. Determine how the probability that a UF student has a Facebbok account changes with the student's age. predict cat. based on quant. → logistic reg.

C 45. Predict the average number of hours per week UF students spend on Facebook, based on the student's age.
1 quant
quant.

Questions 46 – 49 Which drug slows reaction time the most? The following are the reaction times (in milliseconds) for randomly selected subjects who took either Drug A or Drug B.

Drug A		Drug B	
1.96	4	2.11	6
2.24	7	2.43	9
1.71	2	2.07	5
2.41	8	2.71	11
1.62	1	2.50	10
1.93	3	4.84	13
		2.88	12

much bigger = outlier → nonparametric

46. These data represents:
   a) two independent samples, but it would have been better to collect data for matched pairs, since reaction times can vary greatly by individual.
   b) matched pairs, but it would have been better to collect data for two independent samples, since reaction times can vary greatly by individual.
   c) quantitative data, but it would have been better to collect categorical data on whether each subject reacted more slowly under Drug A or B.
   d) categorical data, but it would have been better to collect quantitative data on whether each subject reacted more slowly under Drug A or B.

47. We could analyze this data with a t test or a Nonparametric procedure. When choosing which procedure works best here it's important to note that:
   a) there are more observations for one treatment than the other
   b) the variances of the two groups are quite different
   c) there is an outlier in the data
   d) all of the above
   e) none of the above

48. If we conduct the Wilcoxon Rank-Sum test on this data, the sum of ranks for drug B is:
   a) 25 b) 66 c) 28 d) 21 e) 53 6+9+5+11+10+13+12

49. The best interpretation of the results from the computer output shown below is:
   a) There are significant differences in the mean reaction time for Drug A and B.
   b) There are significant differences in the median reaction time for Drug A and B.
   c) Reaction times are slower for Drug B, on average.
   d) Reaction times are slower for Drug B, overall.

"overall" b/c p-val. significant - would say "on average" if not significant results

50. Are there any problems with the assumptions for the analysis below?
   a) No problems if we can trust the subjects were really random.
   b) No problem since the story states the subjects were chosen randomly.
   c) There is a problem since looking at the data we don't trust the subjects were really random.
   d) There is a problem since the sample size requirement is not satisfied.

No sample size requirement

Mann-Whitney Test and CI: DrugA, DrugB

```
        N  Median
DrugA   6   1.945
DrugB   7   2.500
Point estimate for ETA1-ETA2 is -0.520
96.2 Percent CI for ETA1-ETA2 is (-1.260,-0.110)
W = 25.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0184
```

group A

p-value