Math 159, Spring 2005
Professor Hardin
Exam #2 – sample question solutions

1. (I'm changing the problem: $Y_1 = 2.9$) We observe: $X_1 = 2.1, X_2 = 1.9, X_3 = 2.6, X_4 = 3.3; Y_1 = 2.9, Y_2 = 2.6, Y_3 = 3.7$.

   (a) The idea here is that we only need to count the values that are more extreme than what we saw. Here is a table of ranks:

ranks	1	2	3.5	3.5	5	6	7
our data	X	X	X	Y	Y	X	Y
as extreme	X	X	Y	X	Y	X	Y
more extreme	X	X	X	Y	X	Y	Y
more extreme	X	X	Y	X	X	Y	Y
more extreme	X	X	X	X	Y	Y	Y

   5 out of $\binom{7}{4} = 35$ of our options give test statistics as or more extreme than our data in the lower direction. We also know that we can look at the opposite quantiles by looking at $T' = n(N+1) - T = 19.5$ in our case. That means that any values at or above T' are also "as or more extreme" than our data. (Note that this is only approximate in our case because we have ties.) Thinking from the other end:

ranks	1	2	3.5	3.5	5	6	7
our data	X	X	X	Y	Y	X	Y
more extreme	Y	Y	Y	X	X	X	X
more extreme	Y	Y	X	Y	X	X	X
more extreme	Y	X	Y	Y	X	X	X
more extreme	Y	Y	X	X	Y	X	X

   So, 9 out of 35 = 0.257 of the possible arrangements under H_o are as or more extreme than our data. Our exact p-value is 0.257.

   (b) Give the tails of the exact distribution for this situation (that is, values that are within the lower 10% and upper 10%.)

t	P(T≤ t)
10	1/35 = 0.029
11.5	3/35 = 0.086
12.5	5/35 = 0.143
...	
18.5	30/35 = 0.857
19	32/35 = 0.914
20	33/35 = 0.943
21.5	35/35 = 1

2. c is a reasonable measure of correlation because if there are no ties, the correlation will be +1 if all the pairs are concordant and -1 if all the pairs are discordant. If there are an equal number of concordant and discordant (and the number of ties is zero) then the correlation is zero.

   The number of ties increases the correlation. If we have all ties, the correlation will be 0.5. There is not standard sentiment for what "correlation = 0.5" means in a given situation, and ties (in the sense of $Y_i = Y_j$) indicate a flat regression line and therefore no correlation. We like the idea of adding ties into the correlation coefficient, but it is unclear whether this method is consistent with the understood definition of correlation.

   Looking at pairwise values allows us to see trends that might exist in subgroups and not over an entire range (see HW problem 5.4.2). Additionally, if there are ties in the data, it would be nice to consider them. If we are trying to show a negative correlation, and we are also being conservative, the above correlation will work for us. We like using ranks over using real values because we don't have to worry about outlying values.

3. (a)
   - Out of the 6 pair, 4 are concordant and 2 or discordant. $T = N_c - N_d = 2, \tau = \frac{4-2}{6} = 0.333$, not significant at the 0.1 level. (We could find the exact distribution by keeping X constant and permuting Y.)
   - $T = \sum(R(X_i) - R(Y_i))^2 = 6, \rho = 1 - \frac{6*6}{4(4^2-1)} = 0.4$, again, not significant at the 0.1 level.

   (b)
   - If we are testing whether the slope is significantly different from zero, we use the Spearman's ρ test (above.) We do not have enough evidence to say that the slope is significantly different from zero.
   - To find the slope we need r & s. We can get (from table A11) the quantiles for $T = N_c - N_d$. $\omega_{1-\alpha/2} = \omega_{0.95} = 4$ which gives r = 1, s=6. Not surprisingly (because our sample size is so small), our 90% CI is going to be formed using the largest and smallest slopes. We are 90% confident that the true slope is between (-5,4).

   (c) If in fact the true correlation was around .3 or .4, and we had a bigger sample size, we would see significance in a larger sample. In order for something to big significant, we need to have it be unlikely under the null hypothesis. With sample sizes as small as 4, we can pretty much see anything under the null hypothesis of no correlation. Additionally, a larger sample size would give a shorter confidence interval (regardless of the correlation value.)

4. We can rephrase this data into:

	Rural	Urban	Suburban	row total	ave rank	ave rank sq
0	0	4	1	5	3	9
1	10	11	5	26	18.5	342.25
2	8	1	5	14	38.5	1482.25
3	3	0	1	4	47.5	2256.25
4	1	0	0	1	50	2500
5	0	0	0	0	50.5	2550.25
6	1	0	0	1	51	2601
7	0	0	0	0	51.5	2652.25
8	0	1	0	1	52	2704
R_i	736.5	306	335.5			$\sum t_i \overline{R}_i^2$ =46525

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{c} t_i \overline{R}_i^2 - N(N+1)^2/4\right] = 196.24$$

$$T = \frac{1}{196.24}\left[\frac{736.5^2}{23} + \frac{306^2}{17} + \frac{335.5^2}{12} - \frac{52(53^2)}{4}\right] = 9.96$$

$$\text{p-value} = P(\chi^2_{3-1} \geq 9.96) < 0.01$$

We reject the null hypothesis. It appears as though these groups have different distributions.

The parameterization on page 292 is just a simplification of the calculations given for many ties. Note that it is easy to argue that equation (2) in the original parameterization is identical to equation (9) in the new parameterization.

5. Though the null and alternative hypotheses are written with the parameters p_1 and p_2, because we can consider both of these as binomial situations, the null hypothesis is EXACTLY equivalent to: $H_o : F(x) = G(x)$ (which is what Fisher's exact test is testing.)

   Let X= # of shots that were made out of those players who missed their first shot.
   p-value = $P(X \geq 14) = \sum_{i=14}^{18} \frac{\binom{21}{i}\binom{9}{18-i}}{\binom{30}{18}} = 0.2312$

6. (b) and (d) give exactly the same results. Whenever the permuted test statistic is larger than the observed test statistic in one situation, it's also bigger in the other situation.

   The above is not true for (a) and (c). However, the null distribution is based on the variances being equal (which means the standard deviations are also equal.) The null distribution says that the ratio of either is equal to 1, and the difference between standard deviations for either is equal to 0. If your data are significantly far from one of these hypothesized settings, it should be significantly far from others. Remember, we aren't using assumptions about the data to describe the test statistic, so we don't have to worry that a ratio test would be valid and a difference test wouldn't be. It is likely that your simulated p-values won't be exactly the same, but they should give similar conclusions about significance.

7. (a) A test of a positive slope is the same as testing Spearman's ρ. The data become:

	Golfer							
	1	2	3	4	5	6	7	8
Age (estimated)	4	2.5	5	6	7	8	1	2.5
$R(X_i)^2$	16	6.25	25	36	49	64	1	6.25
Amount Paid	3	5	2	6	7	8	1	4
$R(Y_i)^2$	9	25	4	36	49	64	1	16
$R(X_i) \cdot R(Y_i)$	12	12.5	10	36	49	64	1	10

   $\rho = 0.778$. Using table A10, we see that this correlation is significant at $\alpha = 0.025$. We think, yes, age and payment are correlated.

   (b) We could have used Kendall's ρ or Pearson's correlation coefficient. Kendall's rho will probably give similar results because it is also based on ranks (unless there are unseen patterns in the data such as differences due to gender.) Pearson's correlation coefficient will also likely give a similar correlation (and value of significance) because there are no outliers and the data don't appear to be extremely skewed (though both variables are slightly skewed right.)

8. With ranks the data are:

											Total of Ranks
1st Group	17	4	13.5	11	12	8	15.5	19	15.5	6	121.5
2nd Group	20	5	2	1	9	13.5	7	10	18	3	88.5

   (a) T = 121.5, $\sum_{i=1}^{N} R_i^2 = 2869, T_1 = 1.248$

      i. Use the exact distribution (ignore ties): $P(T \geq 121.5) > 0.1$ (but just barely), p-value ≈ 0.2

      ii. Use the normal approximation: p-value $= 2 \cdot P(T_1 \geq 1.248) = 2 \cdot 0.106 = 0.212$.
      Both methods say that you cannot detect a difference in reading ability between the two methods.

      iii. Both methods give similar p-values. This is not surprising because there are a moderate number of sample points, and there aren't many ties. The exact distribution (if we had calculated it using ties) would probably have given a p-value along the same lines.
      Note: it's a good thing that we were using a nonparametric test here. The values of 19 & 14 were real typos, on my part! (They should have been 193 and 146.)

   (b) To find a 99.8% CI, we need to calculate the $\omega_{0.01}$ cutoff points for the test statistic that compares the pairwise differences. Here, $\omega_{0.01} = 66$. Which gives k = 66 - 10(11)/2 = 11.
   The 11^{th} difference is -95, the 180^{th} difference is 174. We're 99.8% confident that the true difference in reading scores between these two methods is between (-95 points, 174 points.) Because zero is contained in the interval, we cannot make any claims about which method is better.