STA 4502/5507
Introduction to Non-parametric Methods

Sample Exam 2

On my honor, I have neither given nor received unauthorized aid on this examination.

Signature: ______________________________ ***Date:*** __________________

Print Name: ____ ***UFID:*** _________________________

Instructions:

i. This is a 50 minute exam. There are 4 problems, worth a total of 55 points. Maximum score is 50, and whatever point you get above 50 will be carried forward.

ii. You may not use any books, other references, or any other electronic devices during the exam. One page (two-sided) hand-written notes is allowed.

1. In the first 10 days of 2021 pre-season NBA, Atlanta Hawks scored 91,96,99 and Orlando magic scored 97, 86, 100 and 103. These two teams have not played each other so far.

(a) Argue that this can be framed into testing similarity of two independent samples. **(2 pts)**

(b) Note that, the range (maximum and minimum) is quite different for these two teams. In this light, state a proper set of one-sided hypotheses to test with appropriate notation. **(3 pts)**

(c) Obtain the Ansari-Bradley test statistic C. **(5 pts)**

(d) Under the null hypothesis , what is the range of C? **(4 pts)**

(e) Write down the pvalue in terms of $P(C \geq t)$ or $P(C \leq t)$ where t is your answer from part (c). **(3 pts)**

2. Consider the same data from Q1 and note that their means are very similar.

(a) In this light, design a two-sided hypothesis test for testing location shift with proper notation. **(3 pts)**

(b) Why does a two-sided test (as compared to a one-sided one) makes more sense given this data? **(2 pts)**

(c) Compute that Wilcoxon rank-sum statistic W. **(5 pts)**

(d) What sort of values of W should favor H_1? (small, large or both?) **(2 pts)**

3. Again consider the same data from Q1. If we were to test equality of distribution,

(a) Write the appropriate hypotheses with proper notation. **(3 pts)**

(b) Compute the KS test statistic. **(5 pts)**

4. In the context of a two-sided **median equality** test for two groups **of equal size** n, let R_1 and R_2 be the sum of ranks of two groups.

(a) State the appropriate set of hypotheses. **(3 pts)**

(b) Argue why ' Reject if $|R_2 - R_1|$ is large' is an intuitive rejection rule. **(3 pts)**

(c) Show that $|\frac{R_1}{n} - \frac{2n+1}{2}| = |\frac{R_2}{n} - \frac{2n+1}{2}|$. **(5 pts)**
(Hint: What is $R_1 + R_2$?)

(d) Use (c) to show that the Kruskal-Wallis test statistic H can be written as

$$H = \frac{12}{2n+1}\left(\frac{R_2}{n} - \frac{2n+1}{2}\right)^2$$

(7 pts)

Sample Exam 2 Solutions

1. a. Since the two teams did not play each other, the scores don't affect each other. Thus it is okay to assume independence and compare the two samples

b. $H_0: \sigma_1^2 = \sigma_2^2$ vs $H_1: \sigma_1^2 > \sigma_2^2$ where σ_1^2 = variance of OM scores
σ_2^2 = " " AH scores

c. Joint Ranks

86   97   100   103 → OM

$\boxed{91}$ $\boxed{96}$ $\boxed{99}$ → AH

1   2   3   4   3   2   1

$C = 2 + 3 + 3 = 8$

e. under H_1 → C should be high given they share the median the ranks should be in the middle spectrum and thus C should be high

d. Lowest value of C → $1 + 1 + 2 = 4$
Largest " " " → $3 + 4 + 3 = 10$

2. Assumption: $G(t) = F(t + \Delta)$ where G → dist^n fn of AH scores
F → " " " OM "

$H_0: \Delta = 0$
$H_1: \Delta \neq 0$

(b) Since they almost share the mean, there is not enough evidence from data for a specific preference.

(c) Joint ranks:

86   97   100   103 → OM

$\boxed{91}$ $\boxed{92}$ $\boxed{99}$ → AH

1   2   3   4   5   6   7

W = sum of ranks of 2nd sample = $2 + 3 + 5 = 10$

(d) under H_1, small or large values of W mean two-sided equality is somewhat violated and thus favors H_1

3. (a) $H_0: G(t) = F(t)$ for all t
$H_1: G(t) \neq F(t)$ for some t } G, F as before

(b) Please complete.

4. (a) $H_0: \eta_1 = \eta_2$ where η_1 → median of pop^n 1
$H_1: \eta_1 \neq \eta_2$ η_2 → " " " 2

(b) R_1 → sum of ranks of first sample
R_2 → " " " second "

$|R_1 - R_2|$ should be small if both are evenly spread

(c) $R_1 + R_2 = \frac{2n(2n+1)}{2} = n(2n+1)$

$R_1 - \frac{n(2n+1)}{2} = \frac{n(2n+1)}{2} - R_2$

then $\left|\frac{R_1}{n} - \frac{(2n+1)}{2}\right| = \left|\frac{R_2}{n} - \frac{(2n+1)}{2}\right|$

(d) $H = \frac{12}{(2n)(2n+1)} \left\{ n\left(\frac{R_1}{n} - \frac{2n+1}{2}\right)^2 + n\left(\frac{R_2}{n} - \frac{2n+1}{2}\right)^2 \right\}$

equal from part (c)

$= \frac{12}{(2n+1)} \left(\frac{R_2}{n} - \frac{2n+1}{2}\right)^2$