① There are $\binom{6}{3} = 20$ equally likely possibilities to consider. I'll list what the ranks would be in the first sample.

Ranks	W	Ranks	W
123	6	234	9
124	7	235	10
125	8	236	11
126	9	245	11
134	8	246	12
135	9	256	13
136	10	345	12
145	10	346	13
146	11	356	14
156	12	456	15

Null distribution:

W	Prob.
6	.05
7	.05
8	.10
9	.15
10	.15
11	.15
12	.15
13	.10
14	.05
15	.05
	1

For a level-0.20 two-tailed test, we would reject H_0 if

$$\boxed{W \le 7 \text{ or } W \ge 14.}$$

(2) There are $\binom{5}{3} = 10$ possibilities. I will list the ranks for the values from the first sample.

test statistic

Ranks	TS	Ranks	TS
123	0	145	2
124	0	234	0
125	1	235	1
134	0	245	2
135	1	345	3

Null distribution:

TS	Prob.
0	0.4
1	0.3
2	0.2
3	0.1

(3) (a) We test H_0: Listening to statistics talks has no effect on milk yield against H_a: Listening to statistics talks has some effect on milk yield.

I'll do a permutation test. The absolute differences are 2, 4, and 3.

2	3	4	Sum	
+	+	+	9	← observed ← most extreme
+	+	−	1	
+	−	+	3	
+	−	−	5	
−	+	+	5	
−	+	−	−3	
−	−	+	−1	
−	−	−	−9	

The p-value is $2\left(\frac{1}{8}\right) = \frac{1}{4}$.

Since $\frac{1}{4} \le \overset{\alpha}{0.25}$, we $\boxed{\text{reject } H_0}$. At level 0.25, we conclude that listening to statistics talks has an effect on milk yield.

(b) We could have done a sign test or a signed rank test. The p-value would have been the same for these data.

(4)

X	Y
1	300
2	2
3	1

$\Rightarrow$

Rank of X	Rank of Y
1	3
2	2
3	1

Since the ranks are perfectly negatively correlated, $r_s = -1$. However, since the (x, y) points don't fall on a line,

$r > -1$.

(5) (a) I used the Kruskal-Wallis test to test H_0: Mowing height has no effect on phosphorous content ($m_5 = m_{10} = m_{20}$) against H_a: Mowing height has some effect on phosphorous content (m_5, m_{10}, and m_{20} are not all equal). Since the p-value was $0.5671 > 0.10$ (α), I retained H_0. There isn't sufficient evidence to conclude at level 0.10 that mowing height has an effect on phosphorous content.

(b) I could have used a permutation F test.

(6) I must politely disagree. Though the rank sum test doesn't use the raw data, ARE calculations show that under a shift model, the rank sum test is never much less efficient (in terms of power) than the t test, while sometimes being far more efficient. Also, unlike the t test, the rank sum test controls α even for non-normal data.

(7) I did a simulation study in R with 1000 runs. The estimated powers were 0.234 (rank sum), 0.187 (Ansari-Bradley), and 0.134 (K-S). Thus, it appears that the $\boxed{\text{rank sum test}}$ has the best power to detect the difference of interest.

8 a) $H_0: \theta_{.5} = 0.5$ vs $H_a: \theta_{.5} > 0.5$.

I will use the sign test since there is no reason to expect symmetry.

B = # values greater than 0.5 = 6.

$\Rightarrow$ P-value $= P(B \geq 6 \mid B \sim Bin(8, \frac{1}{2}))$

$= \binom{8}{6}\left(\frac{1}{2}\right)^8 + \binom{8}{7}\left(\frac{1}{2}\right)^8 + \binom{8}{8}\left(\frac{1}{2}\right)^8 \cong \boxed{0.145}$.

Since $0.145 > 0.05 \overset{\leftarrow \alpha}{}$, I $\boxed{\text{retain } H_0}$. There is not enough evidence to conclude at level 0.05 that the median exceeds 0.5.

b) Since $n = 8$, the coverage for the prediction interval $(X_{(r)}, X_{(s)})$ $(1 \leq r < s \leq n)$ is

$$\frac{s-r}{n+1} = \frac{s-r}{9}.$$

To get at least 75% coverage, we need $s - r \geq 0.75(9) = 6.75$

$\Rightarrow$ Only $(X_{(1)}, X_{(8)}) = \boxed{(.045, .980)}$ will give a two-sided interval. The exact coverage probability is $\frac{7}{9} \cong \boxed{77.8\%}$