Normal approximations with Wilcoxon tests

Starter

1. **(Review of last lesson)** Eleven randomly selected primary school children are weighed. The results, in kg, are shown in the table below.

Girls	43.7	34.4	53.1	38.3	48.5	48.9
Boys	41.4	36.6	34.6	32.5	32.4	

Test at the $1\ \%$ level, whether the data supports the view that, on average, primary school girls have greater weights than the boys.

Working: H_0 : on average boys and girls have the same weight in primary school
H_1 : on average girls weight more than boys in primary school
Here are the weights ranked in order from smallest to largest:

Value	32.4	32.5	34.4	34.6	36.6	38.3	41.4	43.7	48.5	48.9	53.1
Rank	1	2	3	4	5	6	7	8	9	10	11
Sex	B	B	G	B	B	G	B	G	G	G	G

$R_B = 19$
$m = 5, n = 6$: $\quad m(m+n+1) - R_m = 5(5+6+1) - 19 = 41$
$\therefore W = 19$
The $1\ \%$ critical value for a one-tailed test when $m = 5$, $n = 6$ is 17.
Since $W = 19 \nleq 17$, we do not reject H_0.
There is no evidence to suggest that on average girls weigh more than boys in primary school.

E.g. 1 When a Wilcoxon signed-rank is carried out on a sample size of 30, the value obtained for T is 154. Test whether this result is significant at the $5\ \%$ level for a one-tail test.

Working: Mean $= \dfrac{1}{4}n(n+1) = \dfrac{1}{4} \times 30 \times 31 = 232.5$
Variance $= \dfrac{1}{24}n(n+1)(2n+1) = \dfrac{1}{24} \times 30 \times 31 \times 61 = 2363.75$
$\sigma = \sqrt{2363.75}$
Continuity correction: since $154 < 232.5$, use $154 + 0.5 = 154.5$
Using a calculator:
Normal CD
Lower: -99999
Upper: 154.5
σ: $\sqrt{2363.75}(\approx 48.618)$
μ: 232.5
$p = P(T \leq 154) \approx P(T' \leq 154.5) \approx 0.0543$
Since $p \approx 0.0543 \nleq 0.05$, we do not reject H_0.
The result is not significant at the $5\ \%$ level.

E.g. 2 From a sample size of 52, the values of W_+ and W_- are 420 and 958 respectively. Test at the $2\ \%$ level for a two-tail test whether this result is significant.

Working: Mean $= \dfrac{1}{4}n(n+1) = \dfrac{1}{4} \times 30 \times 31 = 689$

Variance $= \dfrac{1}{24}n(n+1)(2n+1) = \dfrac{1}{24} \times 30 \times 31 \times 61 = 12057.5$

$\sigma = \sqrt{12057.5}$

Continuity correction: since $420 < 689$, use $420 + 0.5 = 420.5$

Using a calculator:

Normal CD

Lower: -99999

Upper: 420.5

σ: $\sqrt{12057.5}(\approx 109.8)$

μ: 689

$p = P(T \leq 420) \approx P(T' \leq 420.5) \approx 0.00724$

Since $p \approx 0.00724 \leq 0.01$, we reject H_0.

The result is significant at the $5\ \%$ level.

E.g. 3 Determine the upper and lower $2\ \%$ one-tailed critical values for a Wilcoxon signed-rank test based on a sample size of 76.

Working: Mean $= \dfrac{1}{4}n(n+1) = \dfrac{1}{4} \times 76 \times 77 = 1463$

Variance $= \dfrac{1}{24}n(n+1)(2n+1) = \dfrac{1}{24} \times 76 \times 77 \times 153 = 37306.5$

$\sigma = \sqrt{37306.5} \approx 193.1$

Using a calculator to calculate the lower critical value:

Inverse Normal

Area: 0.02

σ: $\sqrt{37306.5}(\approx 193.14)$

μ: 1463

The value given is 1066.32

Since this value is less than the mean, it would have had 0.5 added to it so $1066 - 0.5 = 1065.82$.

The ***lower critical value*** is therefore 1065

Using a calculator to calculate the upper critical value:

Inverse Normal

Area: 0.98

σ: $\sqrt{37306.5}(\approx 193.14)$

μ: 1463

The value given is 1859.68

Since this value is more than the mean, it would have had 0.5 subtracted from it so $1859.68 + 0.5 = 1860.18$.

The ***upper critical value*** is therefore 1861.

E.g. 4 A Wilcoxon rank-sum test is carried out on samples of size 12 and 17. The R_m value is 210. State the value of the test statistic, W, you are using and test at the $5\,\%$ level for a one-tail test whether this result is significant.

Working:

$m = 12, n = 17$: $\quad m(m+n+1) - R_m = 12(12+17+1) - 210 = 150$

$\therefore W = 150$

Mean $= \dfrac{1}{2}m(m+n+1) = \dfrac{1}{2} \times 12 \times (12+17+1) = 180$

Variance $= \dfrac{1}{12}mn(m+n+1) = \dfrac{1}{12} \times 13 \times 14 \times (13+14+1) = 510$

$\sigma = \sqrt{510}$

Continuity correction: since $150 < 180$, test statistic is $150 + 0.5 = 150.5$

Using a calculator:

Normal CD

Lower:	-99999
Upper:	150.5
σ:	$\sqrt{510}(\approx 22.583)$
μ:	180

$p = P(W \le 150) \approx P(W' \le 150.5) \approx 0.0957$

Since $p \approx 0.0957 \nleq 0.005$, we do not reject H_0.

The result is not significant at the $5\,\%$ level.

Alternatively:

Since $210 > 180$, the test statistic is $210 - 0.5 = 209.5$

Using a calculator:

Normal CD

Lower:	209.5
Upper:	99999
σ:	$\sqrt{510}(\approx 22.583)$
μ:	180

$p = P(W \ge 210) \approx P(W' \ge 209.5) \approx 0.0957$

E.g. 5 When a Wilcoxon rank-sum test was carried out the value obtained was $W = 126$ with $m = 13$ and $n = 14$. Is this result significant at the $1\,\%$ level for a two-tail test?

Working:

Mean $= \dfrac{1}{2}m(m+n+1) = \dfrac{1}{2} \times 13 \times (13+14+1) = 182$

Variance $= \dfrac{1}{12}mn(m+n+1) = \dfrac{1}{12} \times 13 \times 14 \times (13+14+1) = \dfrac{1274}{3} = 424.\dot{6}$

$\sigma = \sqrt{424.\dot{6}}$

Continuity correction: since $126 < 182$, use $R_m = 126 + 0.5 = 126.5$

Using a calculator:

Normal CD

Lower:	-99999
Upper:	126.5
σ:	$\sqrt{424.\dot{6}}(\approx 20.61)$
μ:	182

$p = P(W \le 126) \approx P(T' \le 126.5) \approx 0.00354$

Since $p \approx 0.00354 \le 0.005$, we reject H_0.

The result is significant at the $1\,\%$ level.

E.g. 6 Determine the upper and lower $10\,\%$ one-tailed critical values for a Wilcoxon rank-sum test based on a sample sizes of $m = 10$ and $n = 14$.

Working: Mean $= \dfrac{1}{2}m(m+n+1) = \dfrac{1}{2} \times 10 \times (10 + 14 + 1) = 125$

Variance $= \dfrac{1}{12}mn(m+n+1) = \dfrac{1}{12} \times 10 \times 14 \times (10 + 14 + 1) = \dfrac{765}{3} = 291.\dot{6}$

$\sigma = \sqrt{291.\dot{6}} \approx 193.1$

Using a calculator to calculate the lower critical value:

Inverse Normal

Area: 0.1

σ: $\sqrt{291.\dot{6}}(\approx 17.078)$

μ: 125

The value given is 103.11

Since this value is less than the mean, it would have had 0.5 added to it so $103.11 - 0.5 = 102.61$.

The ***lower critical value*** is therefore 102

Using a calculator to calculate the upper critical value:

Inverse Normal

Area: 0.9

σ: $\sqrt{291.\dot{6}}(\approx 17.078)$

μ: 125

The value given is 146.88

Since this value is more than the mean, it would have had 0.5 subtracted from it so $146.88 + 0.5 = 147.38$.

The ***upper critical value*** is therefore 148.

E.g. 7 Two cold remedies, A and B, are compared by a consumer group. Two random samples of people with colds were given either remedy A or remedy B. The size of the sample for A was 32 and sample size of B was 27. The sum of the ranks of A was 1086. Test at the $5\,\%$ significance level whether there is difference in the ranks of the two remedies.

Working: $m = 27, n = 32$

The sum of the ranks of the smaller sample is needed.

Sum of ranks $= \dfrac{1}{2}n(n+1) = \dfrac{1}{2} \times (27 + 32) \times (27 + 32 + 1) = 1770$

So the sum of the ranks of sample B $= 1770 - 1084 = 686$

Mean $= \dfrac{1}{2}m(m+n+1) = \dfrac{1}{2} \times 27 \times (27 + 32 + 1) = 810$

Variance $= \dfrac{1}{12}mn(m+n+1) = \dfrac{1}{12} \times 27 \times 32 \times (27 + 32 + 1) = 4320$

$H_0:\ m_A = m_B$

$H_1:\ m_A \neq m_B$ where m_A and m_B are the medians of the rankings given to remedy A and B respectively.

This is a two-tailed test.

$R_B \sim \mathrm{N}(810, 4320)$

Continuity correction: as $R_B = 686 < 810 =$ mean, add 0.5.

$p = P(R_B \le 686) \approx P(T' \le 686.5) \approx 0.0301$

Since $p \approx 3.01\,\% > 2.5\,\%$ (two-tailed test), we do not reject H_0.

There is evidence to suggest that the medians of the cold remedies are not different i.e. there is insufficient evidence that there is a difference in the cold remedies effectiveness.

E.g. 8 A school wants to test whether online learning is as effective as normal class teaching. Two groups of randomly chosen students were taught the same course and then took an exam to test their understanding. Group X, of 18 students, was taught by a classroom teacher while Group Y, of 24 students, followed a course online. The sum of the ranks of Group Y in the exam was 444. Test at the $5\,\%$ significance level whether normal teaching methods are better than online learning.

Working: $m = 18, n = 24$

The sum of the ranks of the smaller sample is needed.

Sum of ranks $= \frac{1}{2}n(n+1) = \frac{1}{2} \times (18+24) \times (18+24+1) = 903$

So the sum of the ranks of sample X $= 903 - 444 = 459$

Mean $= \frac{1}{2}m(m+n+1) = \frac{1}{2} \times 18 \times (18+24+1) = 387$

Variance $= \frac{1}{12}mn(m+n+1) = \frac{1}{12} \times 18 \times 24 \times (18+24+1) = 1548$

$H_0 :\ m_X = m_Y$

$H_1 :\ m_X > m_Y$ where m_X and m_Y are the medians of the rankings of Group X and Y respectively.

This is a one-tailed test.

$R_X \sim \mathrm{N}(387,\ 1548)$

Continuity correction: as $R_X = 459 > 387 =$ mean, subtract 0.5.

$p = P(R_X \geq 459) \approx P(T \geq 458.5) \approx 0.0346$

Since $p \approx 3.46\,\% < 5\,\%$, we reject H_0.

There is evidence to suggest that the median of the ranks of the teacher-led lessons is higher than those of the lone course.

Video (password): Single sample tests: normal approximation for large samples

Video (password): Paired sample tests: normal approximation for large samples

Video: Normal approximations with Wilcoxon tests

Solutions to Starter and E.g.s

Exercise

p60 4E Qu 1-4