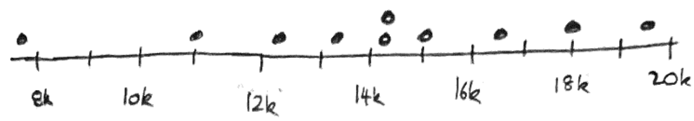
CIMT Further Statistics p53 Ex3E — Solved using Wilcoxon Single Sample test, and not the Sign Test.

1. we have single sample data

we shall assume distribution of annual salaries is symmetrical — this dotplot supports that assumption.



hence do Wilcoxon Signed Rank test

Salaries	median	salaries − median	|salaries − median|	rank
13250	12000	1250	1250	3
7485	12000	−4515	4515	$\boxed{8}$
15136	12000	3136	3136	6
12258	12000	258	258	1
11019	12000	−981	981	$\boxed{2}$
14268	12000	2268	2268	4
19536	12000	7536	7536	10
14326	12000	2326	2326	5
16326	12000	4326	4326	7
17984	12000	5984	5984	9

H_0: salaries have median of £12000

H_1: median salaries > £12000

Assume H_0 to be true

$\alpha = 10\%$

one-tail test

$W_- = 2+8 = 10$

$W_+ = 3+6+1+4+10+5+7+9 = 45$

let $W = \min(W_-, W_+) = 10$.

we want $P(W \le 10)$ when $n = 10$

from tables, $P(W \le 13) = 0.10$
$P(W \le 10) = 0.05$
$P(W \le 8) = 0.25$

so $P(W \le 10) \approx 0.05 < 0.10$.

So, we are inside 10% critical region

We therefore reject H_0 and conclude that there is evidence that population median annual salaries exceed £12000.

2.

A - highest grade
B
C
D
E
F - lowest grade

H_0: population median grade is a C.

H_1: population median grade is not a C.

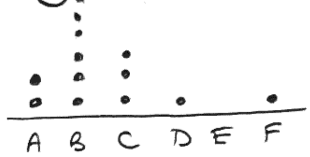
We have single sample data

we assume distribution of grades is symmetrical

(this is doubtful !!)

assume H_0 to be true

$\alpha = 5\%$, two tail test



grade	median	grade − median	|grade − median|	rank
C	C	0		
B	C	1	1	$1 \to 4$
A	C	2	2	$8 \to 8.5$
B	C	1	1	$2 \to 4$
B	C	1	1	$3 \to 4$
C	C	0		
D	C	−1	1	$4 \to \boxed{4}$
C	C	0		
F	C	−3	3	$10 \to \boxed{10}$
B	C	1	1	$5 \to 4$
A	C	2	2	$9 \to 8.5$
B	C	1	1	$6 \to 4$
B	C	1	1	$7 \to 4$

so $W_- = 4 + 10 = 14$

$W_+ = 4 + 8.5 + \ldots + 4 + 4 = 41$

so $W = \min(W_-, W_+)$
$= 14$

we want $P(W \leq 14)$ for $n = 10$

from tables $P(W \leq 13) = 0.10$

$P(W \leq 10) = 0.05$

$P(W \leq 8) = 0.25$

now $P(W \leq 14) > P(W \leq 13) = 0.10$

so we are not in most extreme 2.5% of distribution (2 tail, 5% test)

Hence we do not reject H_0

We conclude that we don't have evidence to suggest that the population median grade is not a C.

3. H_0: population median speed, $\eta = 30$ mph.

H_1: population median speed, $\eta > 30$ mph

Assume H_0 to be true

Assume parent distribution of speeds is symmetrical :

from the stem and leaf plot, this looks doubtful

2	4 9 9
3	5 0 2 5 4 0 8 8 0 4 9 8 0
4	2 2 1 3 0
5	6
6	2
7	2

5|6 = 56 mph

$\alpha = 5\%$.

one tail test

speed	median	speed − median	|speed − median|	rank
24	30	−6	6	8
29	30	−1	1	1 → 1.5
29	30	−1	1	2 → 1.5
30	30	0		
30	30	0		
30	30	0		
30	30	0		
32	30	2	2	3
34	30	4	4	4 → 4.5
34	30	4	4	5 → 4.5
35	30	5	5	6 → 6.5
35	30	5	5	7 → 6.5
38	30	8	8	9 → 10
38	30	8	8	10 → 10
38	30	8	8	11 → 10
39	30	9	9	12
40	30	10	10	13
41	30	11	11	14
42	30	12	12	15 → 15.5
42	30	12	12	16 → 15.5
43	30	13	13	17
56	30	26	26	18
62	30	32	32	19
72.	30	42	42	20

so $W_- = 8 + 1.5 + 1.5 = 11$

$W_+ = 3 + 4.5 + \dots + 19 + 20$

$= 199.$

so $W = \min(W_-, W_+)$

$= 11$

we want $P(W \leq 11)$ for $n = 20$

from tables $n = 20$
$P(W \leq 70) = 0.10$
$P(W \leq 60) = 0.05$
$P(W \leq 52) = 0.25$

so $P(W \leq 11) < 0.05$ and so we are in critical region

we have evidence to reject H_0 and conclude that the population median speed exceeds 30 mph.

4.

adjustments	0	1	2	3	4	5	≥ 6
no. days	61	54	65	18	12	9	31
median	1	1	1	1	1	1	1
adjustments − median	−1	0	1	2	3	4	5+
|adjustments − median|	1		1	2	3	4	5+
rank	$1 \to 61$		$62 \to 126$	127 ↓ 144	145 ↓ 156	157 ↓ 165	166 ↓ 196

$\sum$ days = 250

sum of 126 ranks
$= \frac{1}{2} \times 126 \times 127$
$= 8001$
split equally gives
63·5

H_0: population median adjustments = 1

H_1: population median adjustments $\neq 1$

Assume H_0 to be true

$\alpha = 5\%$

two tail test

Also assume distribution of adjustments is symmetrical (this is doubtful)

so $W_- = 61 \times 63{\cdot}5 = 3873{\cdot}5$

$W_+ = 65 \times 63{\cdot}5 + 127 + \ldots\ldots + 196$

$= 4127{\cdot}5 + 11305$ ← $\frac{1}{2} \times 196 \times 197 - \frac{1}{2} \times 126 \times 127$

$= 15432{\cdot}5$

so $W = \min(W_-, W_+)$

$W = 3873{\cdot}5$

we want $P(W \leq 3873{\cdot}5)$ for $n = 250 - 54 = 196$

we need to do Normal approximation

$E(W) = \frac{1}{4} \times 196 \times 197 = 9653$

$Var(W) = \frac{1}{24} \times 196 \times 197 \times (2 \times 196 + 1) = 632271{\cdot}5$

so $P(W \leq 3873{\cdot}5) \approx P\left(Z < \frac{3873{\cdot}5 - 9653}{\sqrt{632271{\cdot}5}}\right)$

$= P(Z < -7{\cdot}26839)$

≈ 0

so at 5% level, we reject H_0

Hence we conclude that we have evidence that the population median number of adjustments is not 1.

A test for the true <u>mean</u> daily number of adjustments would be difficult as it would require us to assume that the "≥ 6" category were all, say, 6's, when in fact the larger number of adjustments would affect the mean in a way that they do not affect the median.