Stat 411/511

Alternatives to the T-tools

Oct 26 2015

Charlotte Wickham

stat511.cwick.co.nz

Two independent samples

Two sample t-test

Randomization test

Wilcoxon Rank Sum

Today

Doesn't assume Normality and is resistant to outliers

Levene's test

Test for equal population standard deviations

Weds

Welch's t-test

t-test without assumption of equal standard deviations

Two paired samples

Paired t-test

Sign test

Doesn't assume Normality and resistant to outliers, quick

Wilcoxon Signed Rank test

Doesn't assume Normality and resistant to outliers, more efficient

Friday

Wilcoxon Rank Sum

aka Rank Sum Test

aka Mann Whitney Test

Ranks

Instead of using the raw data, we convert the values to ranks.

Ranks are **resistant** to outliers since an outlying value will only ever be 1 unit away from next value.

Removes information about shape (no Normality assumption)

The **test statistic, T,** for the Wilcoxon **Rank Sum** test is the **sum of the ranks in the smaller group.**

Null hypothesis

Assuming the two populations have the same shape and same standard deviation, i.e. the only difference is a shift in center.

Null: The populations distributions are the same. This implies the population **means/medians** are equal (or the difference in population means/medians is zero)

OR

Null: The additive treatment effect is zero

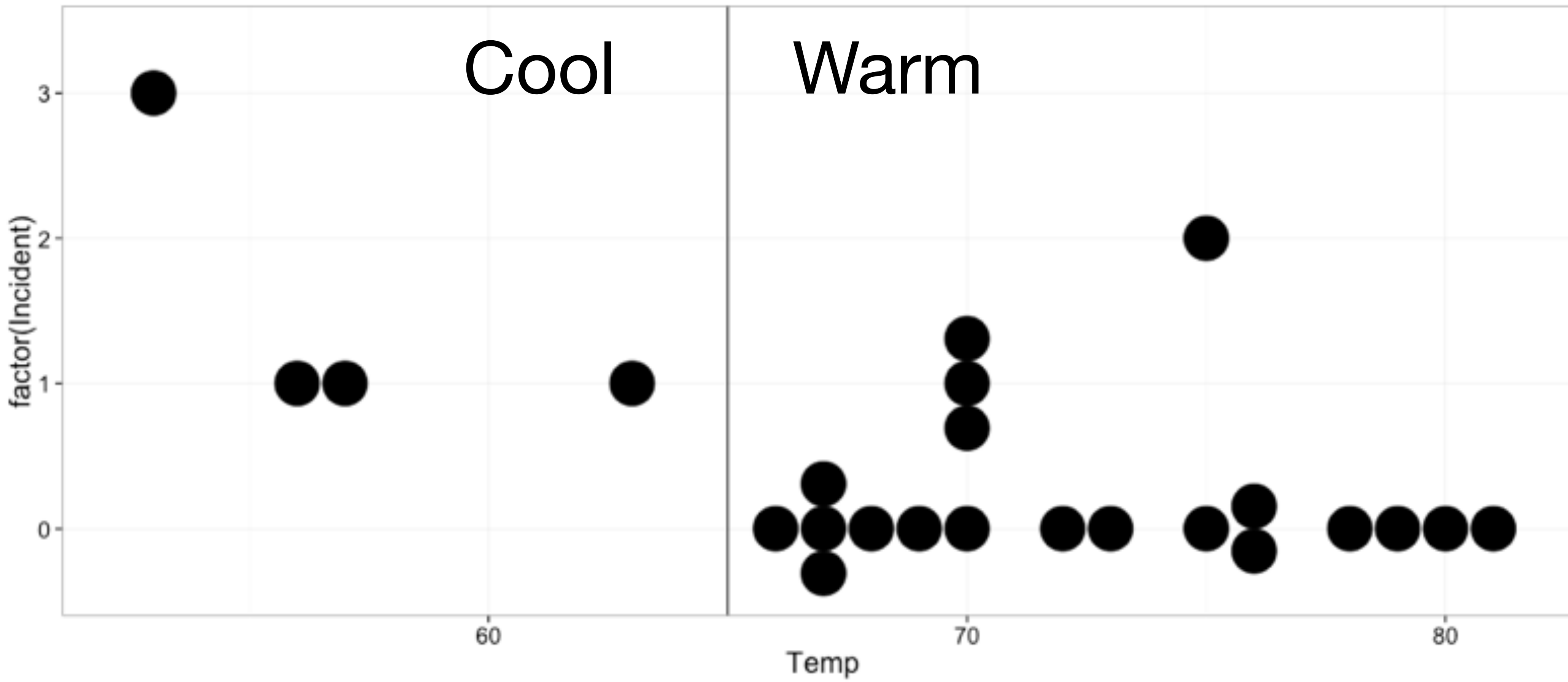
If the two populations don't have the same shape, the test is still valid but it tests,

Null: The population distributions are the same, $P(Y_2 > Y_1) = 0.5$

Alternative: Observations from one population tend to be larger than observations from the other population, $P(Y_2 > Y_1) \neq 0.5$.

case0401

Space shuttle O-ring failures



Are there more O-ring failures at lower temperatures?

Challenger launched on Jan 27 1986 at 29°F

Null: The population mean/median number of O-ring failures at Cool temperatures is the same as at Warm temperatures.

One-sided alternative

Incidents	Launch Temp.
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
1	Warm
1	Warm
2	Warm
1	Cool
1	Cool
1	Cool
3	Cool

Your turn

Step 1:

Using data from both groups, order the observations from smallest response to largest response.

Put the number of incidents in order from smallest to largest, and indicate whether each number is from a Warm or Cool launch

Incidents	Launch
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
0	Warm
1	Cool
1	Cool
1	Cool
1	Warm
1	Warm
2	Warm
3	Cool

Step 2:

Add a new column called "order", with the numbers from 1 to $n_1 + n_2$.

Incidents	Launch	order
0	Warm	1
0	Warm	2
0	Warm	3
0	Warm	4
0	Warm	5
0	Warm	6
0	Warm	7
0	Warm	8
0	Warm	9
0	Warm	10
0	Warm	11
0	Warm	12
0	Warm	13
0	Warm	14
0	Warm	15
0	Warm	16
0	Warm	17
1	Cool	18
1	Cool	19
1	Cool	20
1	Warm	21
1	Warm	22
2	Warm	23
3	Cool	24

Step 2:

Add a new column called "order", with the numbers from 1 to $n_1 + n_2$.

Step 3:

Get ranks by searching for ties, and replacing the order, with the average order for the observations with the same value.

Your turn

Incidents	Launch	order
0	Warm	1
0	Warm	2
0	Warm	3
0	Warm	4
0	Warm	5
0	Warm	6
0	Warm	7
0	Warm	8
0	Warm	9
0	Warm	10
0	Warm	11
0	Warm	12
0	Warm	13
0	Warm	14
0	Warm	15
0	Warm	16
0	Warm	17
1	Cool	18
1	Cool	19
1	Cool	20
1	Warm	21
1	Warm	22
2	Warm	23
3	Cool	24

Replace the order of the observations that are tied, with their average order.

Hint: the average of the integers $m, m + 1, \ldots , n$ is $(m + n)/2$.

```
case0401$rank <- rank(case0401$Incidents)
```

Incidents	Launch	order	rank
0	Warm	1	9
0	Warm	2	9
0	Warm	3	9
0	Warm	4	9
0	Warm	5	9
0	Warm	6	9
0	Warm	7	9
0	Warm	8	9
0	Warm	9	9
0	Warm	10	9
0	Warm	11	9
0	Warm	12	9
0	Warm	13	9
0	Warm	14	9
0	Warm	15	9
0	Warm	16	9
0	Warm	17	9
1	Cool	18	20
1	Cool	19	20
1	Cool	20	20
1	Warm	21	20
1	Warm	22	20
2	Warm	23	23
3	Cool	24	24

Step 4:

Add the ranks of the smallest group

Test statistic,

$T = 20 + 20 + 20 + 24$
$= 84$

just like a randomization test, except since this isn't a randomized experiment, we call it a permutation test.

exact p-value

Elements of smaller group	No. of groupings	Value of T	Value of $\bar{Y}_1 - \bar{Y}_2$
0 0 0 0	2380	36	-0.5
0 0 0 1	3400	47	-0.2
0 0 0 2	680	50	0.1
0 0 1 1	1360	58	0.1
0 0 0 3	680	51	0.4
0 0 1 2	680	61	0.4
0 1 1 1	170	69	0.4
0 0 1 3	680	62	0.7
0 1 1 2	170	72	0.7
1 1 1 1	5	80	0.7
0 0 2 3	136	65	1.0
0 1 1 3	170	73	1.0
1 1 1 2	10	83	1.0
0 1 2 3	85	76	1.3
1 1 1 3	**10**	**84**	**1.3**
1 1 2 3	10	87	1.6
	10626		

List all the ways of reassigning the observed "Incidents" to the "Cool" group.

What proportion have as extreme values of T?

$T \geq 84$

$(10 + 10) / 10626 = 0.0018$

There is convincing evidence the number of O-ring failures at Cool temperatures tends to be larger than the number of failures at Warm temperatures

(Wilcoxon rank-sum, exact one-sided p-value =0.0018).

asymptotic p-value

assume the sampling distribution of test statistic is Normal

Fine if sample sizes are large, and not too many ties.

Under the null the test statistic will:

1. have mean, $n_1\bar{R}$ `= 4 * (1 + 24)/2 = 50`
2. have standard deviation,

   s_R sqrt(($n_1 n_2$)/ (n_1 + n_2))

   `= sd(case0401$rank) * sqrt(4*20/(4 + 20)) = 10.3`
3. be approximately Normal when scaled:
   (T - mean(T)) / sd(T)

```
1- pnorm((84 - 0.5 - 50) / 10.3) = 0.006
```

continuity correction, section 4.2.3

```
wilcox.test(Incidents ~ Launch, data  = case0401,
            alternative = "greater")
    Wilcoxon rank sum test with continuity correction

data:  Incidents by Launch
W = 74, p-value = 0.000572
alternative hypothesis: true location shift is greater
than 0
```

asymptotic

```
Warning message:
In wilcox.test.default(x = c(1, 1, 1, 3), y = c(0, 0, 0,
0, 0, 0,  :
  cannot compute exact p-value with ties
```

```
library(coin)
wilcox_test(Incidents ~ Launch, data  = case0401,
            alternative = "greater", distribution = "exact")

    Exact Wilcoxon Mann-Whitney Rank Sum Test

data:  Incidents by Launch (Cool, Warm)
Z = 3.301, p-value = 0.001882
alternative hypothesis: true mu is greater than 0

> wt <- wilcox_test(Incidents ~ Launch, data  = case0401,
+             alternative = "greater", distribution = "exact")
> statistic(wt, "linear")
Cool 84
```

exact

Your turn

Fill in the column for Wilcoxon Rank Sum on the test summary worksheet (posted on web).