

Causal Inference, Homework 1

Ryan Dew

September 22, 2015

Question 1

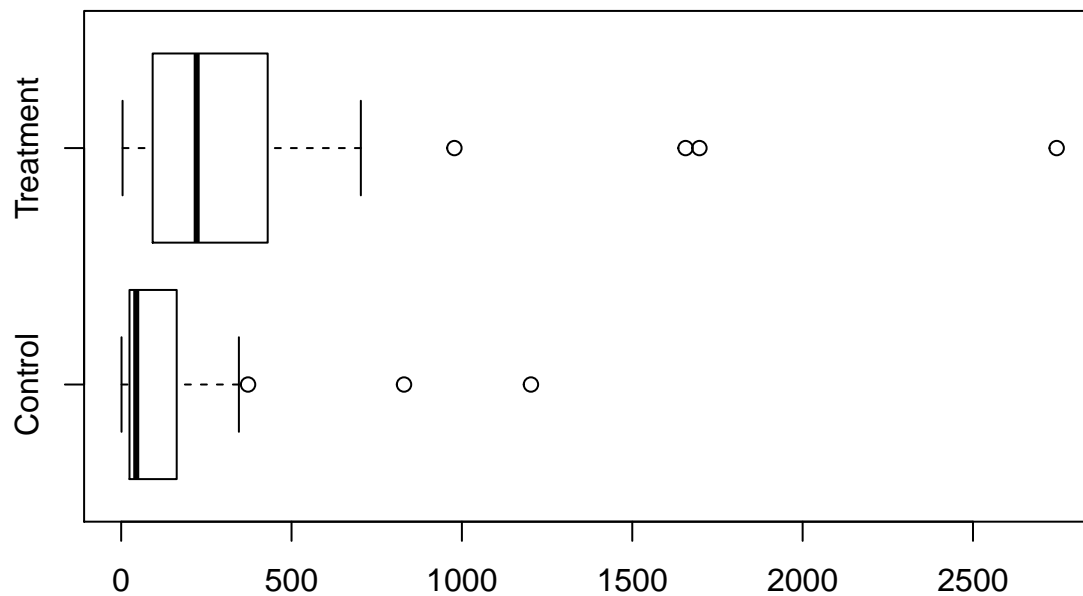
Inputting the data:

```
treatment = c(2745.6, 1697.1, 1656.4, 978, 703.4, 489.1, 430, 334.1, 302.8,  
274.7, 274.7, 255, 242.5, 200.7, 198.6, 129.6, 119, 118.3, 115.3, 92.4, 40.6,  
32.7, 31.4, 17.5, 7.7, 4.1)  
  
control = c(1202.6, 830.1, 372.4, 345.5, 321.2, 244.3, 163, 147.8, 95, 87, 81.2,  
68.5, 47.3, 41.1, 36.6, 29, 28.6, 26.3, 26, 24.4, 21.4, 17.3, 11.5, 4.9, 4.9,  
1.0)
```

(a)

The boxplot is:

```
boxplot(control,treatment,horizontal=T,names=c("Control","Treatment"))
```

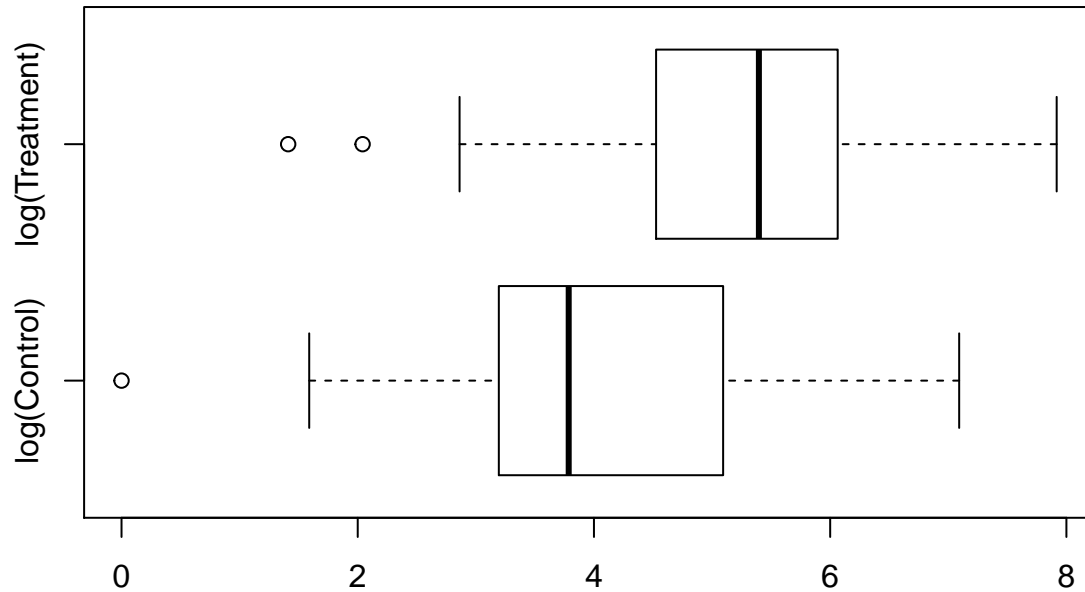


This implies that the additive treatment effect model is NOT reasonable; there appears to be a scaling of the distribution, rather than a shift of the distribution. The variance/interquartile range appears much greater in the treated group than the control.

(b)

The log boxplots are:

```
boxplot(log(control),log(treatment),horizontal=T,names=c("log(Control)","log(Treatment)"))
```



The multiplicative treatment effect seems more reasonable. These two appear to have equal spread, and appear to be simple shifts of each other, as we would expect on the log scale for a multiplicative treatment effects: if $r_{T_i} = \delta r_{C_i}$, then $\log(r_{T_i}) = \log(r_{C_i}) + \log(\delta)$.

(c) and (d)

Testing whether $\delta > 1$ in the multiplicative model is the same as testing whether there is an additive effect of $\log(\delta) = \log(1) = 0$ on the log-scale:

```
rst <- wilcox.test(log(treatment),log(control),alternative = "greater", conf.int = TRUE)
rst
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: log(treatment) and log(control)
## W = 473, p-value = 0.006916
## alternative hypothesis: true location shift is greater than 0
## 95 percent confidence interval:
##  0.3822671      Inf
## sample estimates:
## difference in location
##                1.26038
```

We find a P-value of 0.0069 in favor of rejecting in favor of the alternative hypothesis that there is a positive multiplicative treatment effect greater than $\delta = 1$. The 95% interval for the treatment effect is $[1.4656, \infty]$.

(e)

No, the rank sum test relies on the assumption of independent observations; serial correlation breaks this assumption, and hence the test is no longer valid.

Question 2

The null hypothesis is that the lady has no power of identification—she picks the same arrangement every time. The alternative hypothesis is that she has power of identification; she can perfectly identify the cups every time regardless of the random assignment. Power is the probability of rejecting a false null hypothesis, given the effect size. Here, assuming she can identify the cups, she will with probability one identify them all, which has a chance of $1/70=0.0143 < 0.05$ probability of happening, so we will reject with probability 1. Hence, the power is 1.

Question 3

Inputting the data:

```
marijuana = c(15, 25, 0, 0, 4, 2, 1, 4, 9, 0, 22, 11, 0, 0, 0)
placebo = c(23, 50, 0, 99, 31, 21, 79, 113, 53, 0, 61, 18, 12, 6, 5)
```

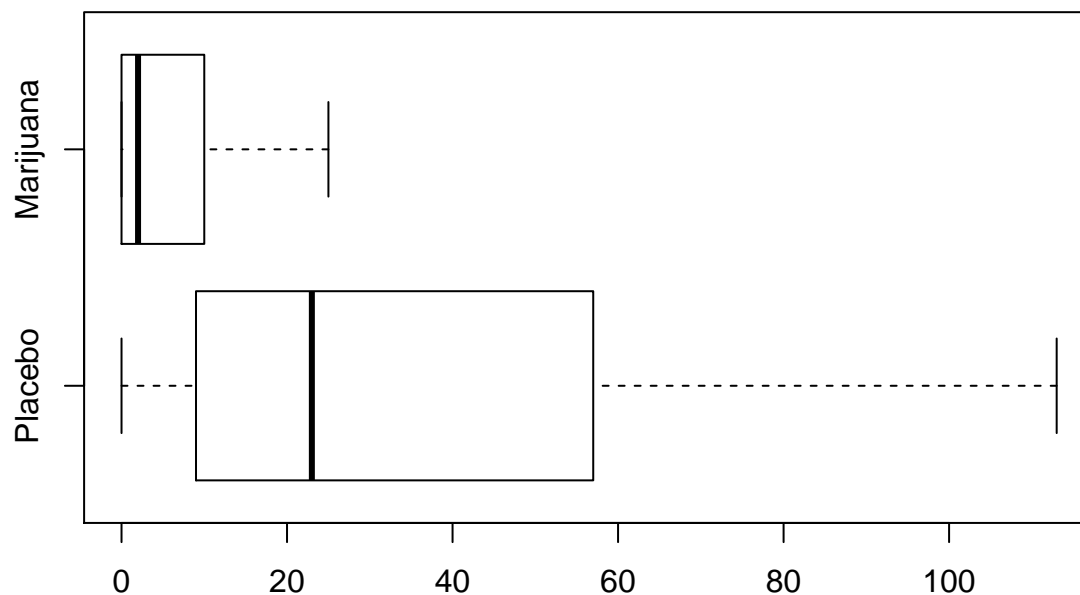
(a)

If this data were available, we could assess several things: first, we could see if the placebo values are different for the group that received it first (and thus had no prior marijuana exposure) than for the group that received it after exposure. We could also see in general if there are any detectable differences between the differenced values of the order groups (so the difference in values for the placebo/marijuana group versus the marijuana/placebo group).

(b)

The additive model is incorrect due to the vastly different spreads apparent in the two boxplots:

```
boxplot(placebo,marijuana,names=c("Placebo","Marijuana"),horizontal=T)
```

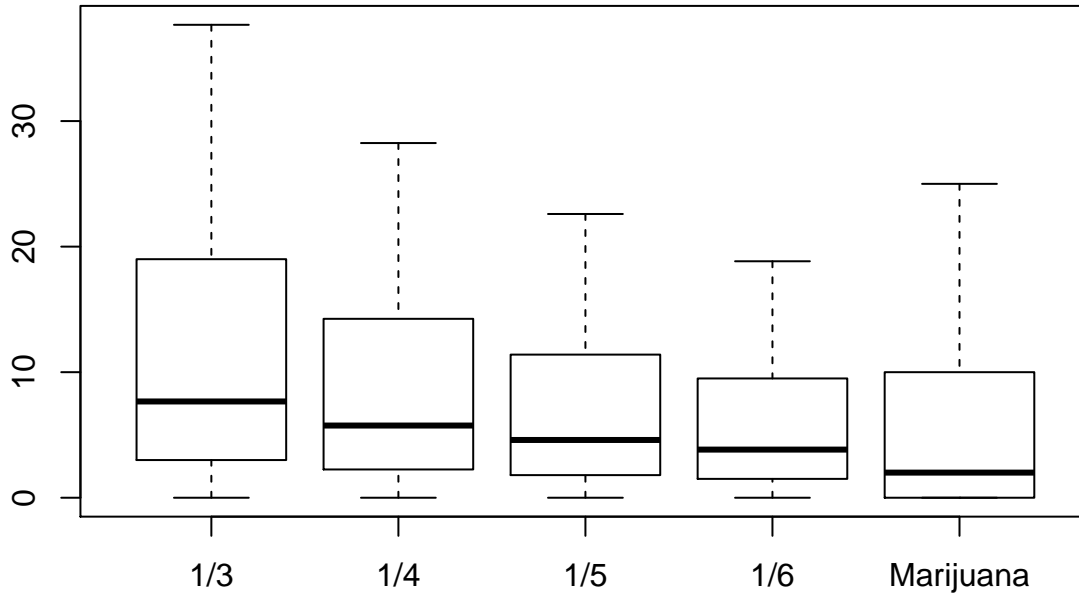


Typically, we would assess a multiplicative effect by looking on the log scale. However, there are a number of zeroes in the data which prevents that. However, we can still assert that a multiplicative treatment model

may not be reasonable because of all these zeroes: if the effect of marijuana reduced the number of vomiting episodes multiplicatively, we would still not expect to see a zero if there were no zeroes in the placebo group. Instead, we see many zeroes. Zeroes are also very meaningful in this context, which makes the multiplicative model even less desirable.

To provide further evidence, I plot here several boxplots pertaining to the multiplicative treatment effects $1/3$, $1/4$, $1/5$, and $1/6$:

```
boxplot(placebo/3,placebo/4,placebo/5,placebo/6,marijuana,
        names=c("1/3","1/4","1/5","1/6","Marijuana"))
```



We see that none appears to be reasonable, although by effect $1/6$, the scale has shrunk past that of the treated group. We see the difference in the distributions is again due to an inability to capture the meaningful zeroes.

(c)

This case is in some sense the mirror image of the case we saw in class, as it still maintains the property that order statistics are preserved. The effect is described as:

$$r_{T_i} = r_{C_i} - \Delta(r_{T_i})$$

for a non-negative, non-decreasing $\Delta(r_{T_i})$. Suppose then that person i has the k th order statistic in the treated outcome, and person j has the $k + 1$ th order statistic for the treated outcome. By definition we have $r_{T_{(k)}} \leq r_{T_{(k+1)}}$ and thus $\Delta(r_{T_{(k)}}) \leq \Delta(r_{T_{(k+1)}})$ by the non-decreasing assumption. Rearranging the above equation, we then have:

$$r_{C_i} = r_{T_i} + \Delta(r_{T_i})$$

Given the two inequalities above, this means that $r_{C_i} \leq r_{C_j}$. This result shows that ranking is preserved from treatment outcomes to control outcomes, or that person i has $r_{C_{(k)}}$.

Since order statistics are preserved, we can approach inference in a similar way as before. Let $\rho = r_{T_{(k)}}$. Then, if observation i corresponds to order k , meaning $r_{T_i} = r_{T_{(k)}}$, we have:

$$R_i - (1 - Z_i)\Delta(\rho) - \rho = r_{T_i} - \rho = r_{C_i} - \Delta(\rho) - \rho = 0$$

If observation i corresponds to order $\ell > k$, then $\Delta(\rho) < \Delta(r_{T_\ell})$; this means, if we adjust our observations by $\Delta(\rho)$, we will not sufficiently adjust i 's outcome: $r_{T_i} - \rho > 0$ and $r_{C_i} - \Delta(\rho) - \rho > r_{C_i} - \Delta(r_{T_i}) - \rho > 0$.

Finally, if observation i corresponds to order $j < k$, then $\Delta(\rho) > \Delta(r_j)$; this means if we adjust our observations by $\Delta(\rho)$, we will over adjust i 's outcome: $r_{T(j)} - \rho < 0$ and $r_{C_i} - \Delta(\rho) - \rho < r_{C_i} - \Delta(r_{T_i}) - \rho < 0$.

In sum, we have:

$$\text{sign}(R_i - (1 - Z_i)\Delta(\rho) - \rho) = \text{sign}(r_{T_i} - \rho)$$

Thus, for testing $H_0 : \Delta(r_{T(k)}) = \Delta_0$ versus the two-sided alternative, we can calculate the adjusted responses $A_j = R_j - (1 - Z_j)\Delta_0$, where now the adjustment is only subtracted if the person was in the control group, and order them $A_{(1)} \leq \dots \leq A_{(N)}$. For simplicity, let $\rho = r_{T(k)}$. Let $q_i = 1$ if $A_i \geq A_{(k)}$, which under the null is equivalent to $q_i = 1$ if $r_{T_i} \geq \rho$. Let $q_i = 0$ otherwise. Then the test statistic given by:

$$T(\Delta_0) = \sum_{i=1}^N q_i(1 - Z_i)$$

which is the number of control subjects whose outcomes under treatment would have met or exceeded ρ . Since this is free of hidden bias due to the matching, $T(\Delta_0)$ has a hypergeometric distribution:

$$\Pr(T(\Delta_0) = t) = \frac{\binom{q_+}{t} \binom{N-q_+}{m-t}}{\binom{N}{m}}$$

where $q_+ = \sum_{i=1}^N q_i$ is the total number of subjects with adjusted outcomes greater than ρ , and m is the number of control subjects, $t = 1, \dots, m$. This can be used for hypothesis testing.

(d)

A 95% confidence interval is the set of null hypotheses that would not be rejected in a 5% hypothesis test. Hence, we write an R function that, for a given k , computes at 95% CI by inverting the hypothesis test described above for $H_0 : \Delta(r_{T(k)}) = \Delta_0$. Note that Δ_0 is restricted to be non-negative, so only values in the positive reals are tested. Also, since all of the data is integer valued, setting Δ_0 to a non-integer does not change any of the orderings, so I check only on an integer-spaced grid.

```
compute.dilated.CI <- function(pairs,k,level=0.95,grid=c(0,1000))
{
  N <- length(pairs)
  if(k>N){stop("k cannot be greater than N")}
  m <- nrow(pairs)
  alpha <- (1-level)/2

  control <- pairs[,1]
  names(control) <- rep("c",length(control))
  treated <- pairs[,2]
  names(treated) <- rep("t",length(treated))

  grid <- seq(grid[1],grid[2],by=1)
  interval <- c()
  for(delta0 in grid)
  {
    A_control <- control-delta0
    A <- sort(c(A_control,treated))

    q <- as.numeric(A>=A[k])
```

```

qplus <- sum(q)

T_stat <- sum(q[names(A)=="c"])

if(phyper(T_stat,qplus,N-qplus,m) > alpha &
    (1-phyper(T_stat-1,qplus,N-qplus,m)) > alpha)
{
  interval <- c(interval,delta0)
}
}
ci <- c(min(interval),max(interval))
ci
}

```

To test that this is working appropriately, I generated some fake data under a dilated treatment model of this type. The output is contained in the appendix, but I am able to capture the truth with the CI every time, supporting that this procedure indeed works.

Now, let's apply this to our data:

```

pairs <- cbind(placebo,marijuana)
ci8 <- compute.dilated.CI(pairs,k=8)
ci15 <- compute.dilated.CI(pairs,k=15)
ci23 <- compute.dilated.CI(pairs,k=23)

```

So the confidence interval for $k = 8$ is given by $[0,12]$, for $k = 15$ is given by $[0,30]$, and for $k = 23$ is given by $[9,95]$.

Appendix

This tests the CI function from question 3:

```

Dfunc <- function(x){10+10*sqrt(x+11)}
fake_pottreat <- runif(100,-10,10)
fake_potctrl <- fake_pottreat+Dfunc(fake_pottreat)
fake_potouts <- cbind(fake_potctrl,fake_pottreat)

compute.dilated.CI(fake_potouts,k=100)

```

```
## [1] 39 44
```

```
Dfunc(sort(fake_pottreat)[50])
```

```
## [1] 41.7128
```

```
compute.dilated.CI(fake_potouts,k=150)
```

```
## [1] 43 52
```

```
Dfunc(sort(fake_pottreat)[75])
```

```
## [1] 47.65001
```

```
compute.dilated.CI(fake_potouts,k=50)
```

```
## [1] 28 38
```

```
Dfunc(sort(fake_pottreat)[25])
```

```
## [1] 33.49678
```

We see the truth is captured in the interval each time.