STA104: Take Home Project 1

Analysis of hours of pain relief provided by two analgesic drugs in 12 patients suffering from arthritis.

Kevin Xu (913288760)

kivxu@ucdavis.edu

Amy T. Kim Instructor

STA104: Nonparametric Statistics

University of California, Davis

Jan. 31, 2021

I.Introduction

The following paper addresses the question of whether an over-the-counter analgesic drug (Drug A or Drug B) provides more hours of pain relief (in hours) for patients suffering from arthritis. Arthritis is a common medical condition that afflicts a large number of elderly people throughout the United States. The results of the paper have implications on medical prescriptions allocated by doctors for severity of arthritis pain, or for specific patient-knowledge so the patient will not be heavily affected by pain for a planned period of time.

My paper will base its analysis in calculating the permutation distribution of differences in median pain relief period between groups 1 and 2 (corresponding to distributions of pain relief period associated with drugs A and B). I will use a permutation test to find the approximate p-value, and confidence interval for the exact p-value. I will then compare p-values of the observed test statistic from sample data under Permutation Test (with Medians), Wilcoxon Rank-Sum test, and Kolmogorov Test to draw inferences of the whether there is a difference in pain relief duration between Drug A and B.

II. Summary of Data

The paper utilizes pain relief duration data (in hours) from subjects' use of Drug A and Drug B of a sample of 12 patients that suffer from arthritis, originating from STA104's "Drug-1" dataset.

Statistical Summary					
	Sample Mean	Sample Standard Deviation	Five Number Summary (Minimum, Quartile 1, Median, Quartile 3, Maximum)	Sample Size
Drug A	6.75833	3.014649	{1.40, 5.25, 6.20, 8.10, 13.70}	12
Drug B	4.69167	1.236166	{2.20, 4.15, 4.50, 5.15, 7.20}	12

Figure 1: Summary statistics of Sample Mean, Standard Deviation, Five Number Summary, and Sample Size of Group A and B.

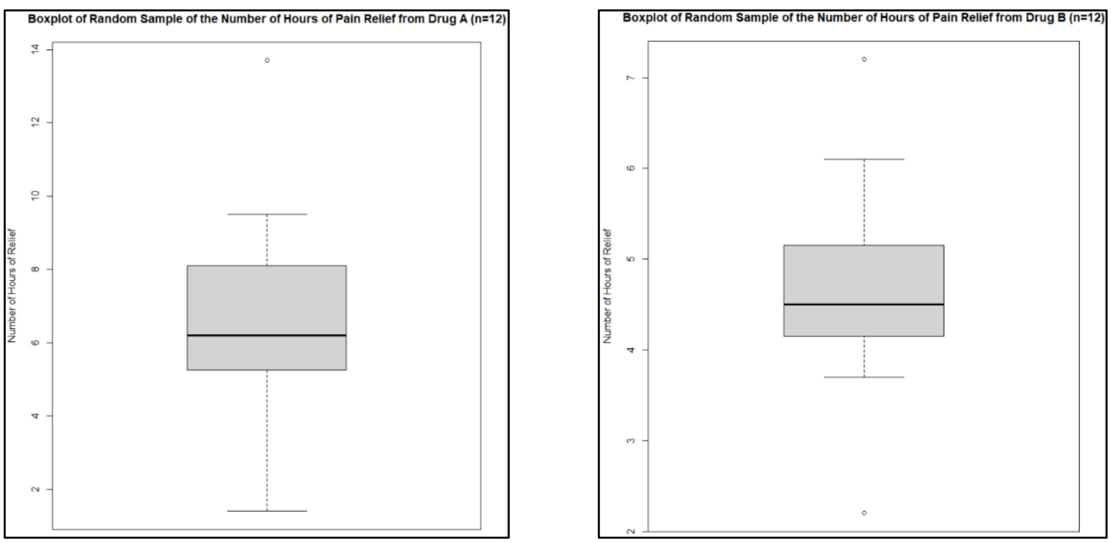

Figure 2: Box Plots indicate that the distribution of sample data from Drug A is skewed left and distribution of sample data from Drug B is skewed right: more data points exist below the 25^{th} percentile for "Drug A group", more data points exist greater than 75^{th} percentile for "Drug B group".

Given that the sample data is skewed for both drug group (asymmetric distribution), the permutation test with differences in median pain relief hours will be used. Since the calculation of medians is not a function of the sum of all pain relief hours within a drug group, it is not as affected by outliers as differences in mean pain relief hours or differences in total pain relief hours between drug groups. Hence, differences in medians is chosen as the method in which the permutation test is conducted. Using this test improves the robustness of our analysis. [Code for Summary of Data is shown in *Code 1* in the Appendix]

III. Analysis

The problem at hand (whether or not pain relief duration is the same between the drug A and drug B) can be equivalently interpreted using the cumulative distribution functions of pain relief duration associated with Drug A and Drug B.

- Group 1 contains sample data for pain relief durations of patients who were assigned Drug A. Group 1 has CDF: $F_1(x)$.
- Group 2 contains sample data for pain relief durations of subjects who were assigned Drug B. Group 2 has CDF: $F_2(x)$.

Firstly, I construct the hypotheses of the problem at hand using a two-tail test:

$$H_0: F_1(x) = F_2(x), \text{ vs, } H_a: F_1(x) \leq F_2(x) \text{ } or \text{ } F_1(x) \geq F_2(x).$$

The two-tail test is selected since the problem does not specify to test if Drug A or Drug B possesses a higher pain relief duration, but rather if there is a difference in pain relief duration. $H_0: F_1(x) = F_2(x)$ is used to assume that selecting data from either group has equal probability.

Secondly, I find the observed test statistic D_{Obs}:

Note:

- $D_i: differences\ in\ medians\ from\ Group\ 1\ and\ Group\ 2\ for\ i^{th} permutation,$
- $D_{Obs}: differences\ of\ medians\ from\ Group\ 1\ and\ Group\ 2\ from\ observed\ sample\ data.$

$$D_{Obs} = Median_1 - Median_2 = 6.20 - 4.50 = 1.70$$

Thirdly, I calculate the P-value of observing test statistic or more extreme:

Since there are $\binom{24}{12} = 2704156$ possible D_i values of the data at hand, it is difficult to calculate the exact p-value of the permutation test. Instead, I find the approximate P-value with the approximate permutation test: which involves randomly generating R permutations. I have chosen to generate $R = 3{,}000$ permutations. Based on R randomly generated D_i values, we have an approximate permutation distribution. *Note, as R increases, the p-value will approach the true P-value.*

Approximate P-value: $\hat{p} = 0.009$ [Code for Approximated P-value is shown in *Code 2*]

Fourthly, I estimate the true p-value of observing our observed statistic ($D_{Obs} = 1.70$) using the binomial distribution. This follows since the approximated p-value $\hat{p}$ is approximately distributed as $N\left(p, \sqrt{\frac{\hat{p}(1-\hat{p})}{R}}\right)$.

CI for the True P-value for D_Obs = 1.70 ($\hat{p}$=0.009)		
α	Lower P-value for D_Obs = 1.70	Upper P-value for D_Obs = 1.70
0.1	0.006163879	0.01183612
0.05	0.005620554	0.01237945
0.01	0.004558655	0.01344135

Figure 3: The 90%, 95%, and 99% confidence intervals for true p-value of $D_{Obs} = 1.70$ using approximated p-value of $\hat{p} = 0.009$ based on R=3000 randomly generated permutations. Code shown in *Code 3*

Lastly, I compared the confidence intervals for true P-value under the Permutation test (median difference between groups) with the Wilcoxon Rank-Sum test P-value and Kolmogorov-Smirnov test P-value to identify if there is a difference between pain relief duration between Drug A and Drug B:

	Test Statistic:	P-value:
Permutation Test (Difference in Medians)	$D_{Obs} = Median_1 - Median_2 = 1.70$	$0.00562 \leq Pvalue \leq 0.0124$ ($95\%\ CI\ for\ true\ P - value$)
Wilcoxon Rank Sum Test	$W_{Obs} = Sum\ of\ Ranks\ in\ Group\ 1 = 194.5$	0.008683
Kolmogorov-Smirnov Test	$K_S = Maximum\ \lvert\widehat{F_1}(x) - \widehat{F_2}(x)\rvert = 0.58333$	≈ 0.033702

Figure 4: Comparison of Test Statistics and P-values between Permutation test (Median difference), Wilcoxon Rank-Sum Test, and Kolmogorov-Smirnov Test. Note: K.S. test P-value is approximate since there are ties present in the sample data. [Code shown in *Code 4*.]

P-values for the Permutation Test, Wilcoxon Rank-Sum Test, and Kolmogorov-Smirnov Test all appear to be significant at a 95% confidence level.

Ⅳ. Interpretation

I take $\alpha = 0.05$. Due to the significant 95% confidence interval ($0.00562 \leq True\ Pvalue \leq 0.0124$) of the Permutation Test, in conjunction with the significant Wilcoxon Rank-Sum Test P-value (0.008683) and significant Kolmogorov-Smirnov Test Approximate P-value (≈ 0.033702), under the null hypothesis that the cumulative distribution function of pain relief duration (in hours) for subjects who took Drug A is the same as that of subjects who took Drug B, I must reject the null hypothesis that the cumulative distribution function associated with the pain relief duration of consuming either drug, is equal. I conclude that the cumulative distribution functions associated with pain relief duration of Drug A and Drug B, are not the same.

Additionally, I am 95% confident that the true pain relief duration of Drug A and Drug B are not the same, and that there is evidence that one analgesic over-the-counter drug (Drug A or Drug B) provides longer pain relief than the other.

Ⅴ. Conclusion

In this analysis, based on the box-plot of pain relief duration for Drug A and B, I have interpreted the distribution of pain relief duration associated with Drug A and Drug B to be skewed (asymmetric). Hence, I conducted tests using the Permutation test (differences in group median), Wilcoxon Rank-Sum test due to their robustness, and because they are less affected by outlier observations. These tests were compared with the Kolmogorov-Smirnov Test.

Furthermore, I calculated the observed/test statistic: a difference in 1.70 hours in pain relief between the medians of group 1 and group 2 (associated with subjects who consumed drug A and drug B respectively). Next, I calculated approximate p-value (0.009) of the observed statistic (1.70 hours) based on an approximated permutation distribution that utilized $R = 3000$ randomly generated permutations. An approximated permutation distribution is used due to the large number of permutations ($24C12 = 2704156$) that would have been needed to be calculated for the exact P-value. I estimated the exact P-value for the observed statistic using the binomial distribution with 90%, 95%, and 99% confidence intervals; it allowed my approximated p-value to be approximately normally distributed. Lastly, I compared the P-values of the Wilcoxon Test, and Kolmogorov-Smirnov Test to that of the Permutation Test with Differences in Medians.

Based on the significant P-values under the Permutation test, Wilcoxon Rank-Sum test, and Kolmogorov test, I reject the null hypothesis that Drug A and B provide the same pain relief duration for arthritis afflicted patients, and I conclude that I am 95% confident that the pain relief duration for arthritis afflicted subjects that consume analgesic over-the-counter drugs, Drug A or Drug B, is not the same.

Inferences gathered here allow me to conclude that doctors should not prescribe Drug A and Drug B interchangeably to arthritis patients due to the difference in pain relief duration associated with Drug A and Drug B. Additionally, arthritis afflicted patients should not use Drug A or B interchangeably in situations that desired pain relief duration is not identical.

Appendix

Code1:

```
> #Summary Statistics:
> drugArows <- Drug.1[Drug.1[,3] == "A",]
> drugArelief <- drugArows$Relief
> drugBrows <- Drug.1[Drug.1[,3] == "B",]
> drugBrelief <- drugBrows$Relief
> drugAmean = mean(drugArelief)
> drugBmean = mean(drugBrelief)
> drugAmean
[1] 6.758333
> drugBmean
[1] 4.691667
> drugAsd = sd(drugArelief)
> drugBsd = sd(drugBrelief)
> drugAsd
[1] 3.014649
> drugBsd
[1] 1.236166
> fivenumA = fivenum(drugArelief)
> fivenumB = fivenum(drugBrelief)
> fivenumA
[1]  1.40  5.25  6.20  8.10 13.70
> fivenumB
[1] 2.20 4.15 4.50 5.15 7.20
> Asamplesize = nrow(drugArows)
> Bsamplesize = nrow(drugBrows)
> Asamplesize
[1] 12
> Bsamplesize
[1] 12
> hist(drugArelief, xlab="Number of hours of relief", ylab="Number of Subjects", main="Histogram of Random Sample of the Number of Hours of Pain Relief from D
rug A (n=12)")
> hist(drugBrelief, xlab="Number of hours of relief", ylab="Number of Subjects", main="Histogram of Random Sample of the Number of Hours of Pain Relief from D
rug B (n=12)")
> boxplot(drugArelief, ylab="Number of Hours of Relief", main="Boxplot of Random Sample of the Number of Hours of Pain Relief from Drug A (n=12)")
> boxplot(drugBrelief, ylab="Number of Hours of Relief", main="Boxplot of Random Sample of the Number of Hours of Pain Relief from Drug B (n=12)")
```

Code 2:

```
> library(coin)
> all.perms = sapply(1:3000,function(i){
+   the.numbers = Drug.1$Relief
+   the.groups = as.factor(Drug.1$Drug)
+   change.groups = sample(the.groups,length(the.groups),replace = FALSE) # shuffles groups
+   group.1.med =  median(the.numbers[change.groups == levels(the.groups)[1]]) # finds median for group 1
+   group.2.med = median(the.numbers[change.groups == levels(the.groups)[2]]) # finds median for group 2
+   difference.in.meds= group.1.med-group.2.med #finds difference in means
+   return(difference.in.meds)
+ })
> difference = 1.70 #finds difference in median
> p.value.two = mean(abs(all.perms) >= abs(difference)) #calculates two-sided p-value
> p.value.two
[1] 0.009
```

Code 3:

```
> p=0.009
> R=3000
> Z_0.95 = qnorm(0.95,mean=0,sd=1,lower.tail = TRUE)
> Z_0.975 = qnorm(0.975,mean=0,sd=1,lower.tail = TRUE)
> Z_0.995 = qnorm(0.995,mean=0,sd=1,lower.tail = TRUE)
> Z_0.95
[1] 1.644854
> Z_0.975
[1] 1.959964
> Z_0.995
[1] 2.575829
> #90% CI for p-value
> a = p - ((Z_0.95) * sqrt(p*(1-p)/R))
> b = p + ((Z_0.95) * sqrt(p*(1-p)/R))
> a
[1] 0.006163879
> b
[1] 0.01183612
> #95% CI for p-value
> a = p - ((Z_0.975) * sqrt(p*(1-p)/R))
> b = p + ((Z_0.975) * sqrt(p*(1-p)/R))
> a
[1] 0.005620554
> b
[1] 0.01237945
> #99% CI for p-value
> a = p - ((Z_0.995) * sqrt(p*(1-p)/R))
> b = p + ((Z_0.995) * sqrt(p*(1-p)/R))
> a
[1] 0.004558655
> b
[1] 0.01344135
```

Code 4:

```
> #Wilcoxon Sum Rank Test:
> sortedDrug = sort(Drug.1$Relief)
> sortedDrug
 [1]  1.4  2.2  3.7  4.1  4.2  4.4  4.5  4.5  5.0  5.1  5.1  5.2  5.2  5.3  5.4  5.7  6.1  6.7  7.0  7.2  7.3  8.9  9.5 13.7
> rankedDrug = rank(sortedDrug)
> rankedDrug
 [1]  1.0  2.0  3.0  4.0  5.0  6.0  7.5  7.5  9.0 10.5 10.5 12.5 12.5 14.0 15.0 16.0 17.0 18.0 19.0 20.0 21.0 22.0 23.0 24.0
> W_obs = 16 + 18 + 9 + 1 + 19 + 14 + 23 + 21 + 15 + 22 + 24 + 12.5
> W_obs
[1] 194.5
> numbers= Drug.1$Relief
> groups = Drug.1$Drug
> trip = data.frame(numbers,groups)
> group = as.factor(groups)
> library(coin)
> wilcox_test(numbers ~ group,data=trip,distribution = "exact",alternative = "two.sided")

        Exact Wilcoxon-Mann-Whitney Test

data:  numbers by group (A, B)
Z = 2.5709, p-value = 0.008683
alternative hypothesis: true mu is not equal to 0
```

```
> split.up = split(Drug.1$Relief,Drug.1$Drug)
> Group1 = split.up[[1]]
> Group2 = split.up[[2]]
> test1 = ks.test(Group1,Group2,alternative = "two.sided",exact = FALSE)
Warning message:
In ks.test(Group1, Group2, alternative = "two.sided", exact = FALSE) :
  p-value will be approximate in the presence of ties
> test1$statistic
        D
0.5833333
> test1$p.value
[1] 0.03370224
```