

# STATS 205: Homework Assignment 5

Brian Liu

6/10/2019

## Solution to Problem 1

We say that two observations  $X_1$  and  $X_2$  are *independent* of one another with respect to a collection of events  $\mathcal{A}$  if

$$Pr \{X_1 \in A \text{ and } X_2 \in B\} = Pr \{X_1 \in A\} Pr \{X_2 \in B\}$$

where A and B are any two not necessarily distinct sets of outcomes belonging to  $\mathcal{A}^3$ .

– 2.2.1 Independent Observations; Permutation, Parametric, and Bootstrap Tests of Hypotheses; Good, Phillip I

In deciding whether your own observations are exchangeable and a permutation test applicable, the key question is the one we posed in the very first chapter: Under the null hypothesis of no differences among the various experimental or survey groups, can we exchange the labels on the observations without significantly affecting the results?

– 2.2.2 Exchangeable Observations; Permutation, Parametric, and Bootstrap Tests of Hypotheses; Good, Phillip I

## Solution to Problem 2

```
cysticerici <- c(28.9, 32.8, 12.0, 9.9, 15.0, 38.0, 12.5, 36.5, 8.6, 26.8);cysticerici
```

```
## [1] 28.9 32.8 12.0 9.9 15.0 38.0 12.5 36.5 8.6 26.8
```

```
worms_reco <- c(1.0, 7.7, 7.3, 7.9, 1.1, 3.5, 18.9, 33.9, 28.6, 25.0); worms_reco
```

```
## [1] 1.0 7.7 7.3 7.9 1.1 3.5 18.9 33.9 28.6 25.0
```

The null hypothesis is that the mean weight of introduced cysticerici *has no correlation with* the mean weight of worms recovered. That is,

$$H_0 : \tau = 0$$

The alternative hypothesis is that the mean weight of introduced cysticerici is *positively correlated with* the mean weight of worms recovered. That is,

$$H_A : \tau > 0$$

To test the null hypothesis against the alternative hypothesis, we will use the Kendall test, a distribution-free test for independence based on signs.

```
cor.test(x = cysticerici, y = worms_reco, method = "kendall", alt = "greater")
```

```
##
```

```
## Kendall's rank correlation tau
```

```
##
```

```
## data: cysticerci and worms_reco
## T = 19, p-value = 0.7578
## alternative hypothesis: true tau is greater than 0
## sample estimates:
##      tau
## -0.1555556
```

The  $p$ -value is 0.7578, which is *not* significant at the  $\alpha = 0.05$  level. There is *not enough* evidence that the mean weight of introduced cysticerci is *positively correlated with* the mean weight of worms recovered.

### Solution to Problem 3

```
cysticerci <- c(28.9, 32.8, 12.0, 9.9, 15.0, 38.0, 12.5, 36.5, 8.6, 26.8)
worms_reco <- c(1.0, 7.7, 7.3, 7.9, 1.1, 3.5, 18.9, 33.9, 28.6, 25.0)
cor.test(x = cysticerci, y = worms_reco, method = "kendall", alt = "greater")
```

```
##
## Kendall's rank correlation tau
##
## data: cysticerci and worms_reco
## T = 19, p-value = 0.7578
## alternative hypothesis: true tau is greater than 0
## sample estimates:
##      tau
## -0.1555556
```

The estimate for  $\tau = -0.1555556$ .

### Solution to Problem 4

### Solution to Problem 5

```
cysticerci <- c(28.9, 32.8, 12.0, 9.9, 15.0, 38.0, 12.5, 36.5, 8.6, 26.8)
worms_reco <- c(1.0, 7.7, 7.3, 7.9, 1.1, 3.5, 18.9, 33.9, 28.6, 25.0)
```

The null hypothesis is that the mean weight of introduced cysticerci *has no correlation with* the mean weight of worms recovered. That is,

$$H_0 : r_s < r_{s,\alpha}$$

The alternative hypothesis is that the mean weight of introduced cysticerci is *positively correlated with* the mean weight of worms recovered. That is,

$$H_A : r_s \geq r_{s,\alpha}$$

Otherwise, do not reject.

To test the null hypothesis against the alternative hypothesis, we will use the Spearman test, a distribution-free test for independence based on ranks.

```
library(SuppDists)
qSpearman(p = 0.05, r = 10)
```

```
## [1] -0.5393939
```

Since  $r_{s,\alpha} = -0.5393939$ , we will reject the null hypothesis only if  $r_s \geq -0.5393939$ .

Calculating  $r_s$ ,

```
cor(x = cysticerici, y = worms_reco, method = "spearman")
```

```
## [1] -0.2
```

Since  $r_s = -0.2$  and  $r_{s,\alpha} = -0.5393939$ , the statement  $r_s \geq r_{s,\alpha}$  is *true*. Thus, we *reject* the null hypothesis. There is *sufficient* evidence that the mean weight of introduced cysticerici is *positively correlated with* the mean weight of worms recovered.

**NOTE:** ~~At this point, I tried to use `cor.test()` with `method = "spearman"` but I got a different result than I expected, and I'm not sure why. Maybe I'm interpreting the output incorrectly?~~

```
cor.test(x = cysticerici, y = worms_reco, method = "spearman", alternative = "greater")
```

```
##
## Spearman's rank correlation rho
##
## data: cysticerici and worms_reco
## S = 198, p-value = 0.72
## alternative hypothesis: true rho is greater than 0
## sample estimates:
## rho
## -0.2
```

~~The  $p$ -value is 0.72, which is *not* significant at the  $\alpha = 0.05$  level. There is *not enough* evidence that the mean weight of introduced cysticerici is *positively correlated with* the mean weight of worms recovered.~~

## Test of methods

```
x = c(7.1, 7.1, 7.2, 8.3, 9.4, 10.5, 11.4)
y = c(2.8, 2.9, 2.8, 2.6, 3.5, 4.6, 5.0)
```

```
qSpearman(0.01, 7)
```

```
## [1] -0.7857143
```

```
cor(x, y, method="spearman")
```

```
## [1] 0.7
```

```
cor.test(x, y, method = "spearman", alternative = "greater")
```

```
## Warning in cor.test.default(x, y, method = "spearman", alternative =
## "greater"): Cannot compute exact p-value with ties
```

```
##
## Spearman's rank correlation rho
##
## data: x and y
## S = 16.8, p-value = 0.03996
## alternative hypothesis: true rho is greater than 0
## sample estimates:
## rho
## 0.7
```

## Solution to Problem 6