Simonson_HW3

Martin A. Simonson

February 8, 2019

1.) Testing insulating fluids

Researchers examined the time *(in minutes)* before an insulating fluid lost its insulating property. The following data are the breakdown of times for 30 samples of the fluid, which had been randomly allocated to receive one of the two voltages of electricity:

- 12 samples assigned to 26kV
- 18 samples assigned to 28kV

Following table gives the summary statistics of the times *(in minutes)* and log-times

Voltages	26kV	28kV	26kV	28kV
.	Time (minutes)	Time (minutes)	log-Time	log-Time
———————-	——————	——————	——————	——————-
Mean	450.89	425.93	5.699	5.620
s.d.	539.86	397.42	0.914	1.098

(a) Is there any evidence that the original responses are not normally distributed? Explain.

- **Answer:** A rule-of-thumb is that a normal distribution should have a standard deviation less than half of that of the mean. The original units of time (minutes) have standard deviations close to or larger than the mean statistic, therefore, there is evidence that the data is not evenly distributed.

(b) Test for a difference between the median breakdown times at two voltages. Clearly write the null and alternate hypothesis, explaining any notation used. Report the **pooled t-statistic and p-value**. Show computations.

Let μ_1 represent the log-transformed mean breakdown time ($\ln(minutes)$) of an insulating fluid at 26kV, and let μ_2 represent the log-transformed mean breakdown time ($\ln(minutes)$) of an insulating fluid at 28kV.

- $H_0 : \mu_1 = \mu_2$
- $H_A : \mu 1 \neq \mu_2$

```
# difference between two means
Mu.diff<-5.699-5.620 # minutes

# pooled sd
Sd.pool<-sqrt((((12-1)*(0.914^2))+(18-1)*(1.098^2)) /
                (12+18-2))

# Standard Error
SE<-Sd.pool * sqrt((1/12)+(1/18))

# T statistic
T.stat<- Mu.diff/SE # with 28 degrees of freedom
T.stat
```

```
## [1] 0.2058764
```

```r
# p-value
p<-2*pt(-abs(T.stat),df=28)
p
```

```
## [1] 0.838377
```

- **Answer:** The pooled t-statistic for this problem is 0.2059 with 28 degrees of freedom. This results in a p-value of 0.8384.

(c) Provide a conclusion in the context of this problem

- **Answer:** There is very weak evidence to reject the null hypothesis. Therefore, we can conclude that the data show that there is no difference in median beakdown times of fluids at the two voltages.

2.) Antibiotic effectiveness

A pharmaceutical company is conducting an experiment to test the effects of a new antibiotic on bacterial counts in tissue cultures. Forty randomly selected tissue cultures were laced in vials and infected with a low quantity of bacterial cells. Of those, 25 were randomly selected and given a dose of antibiotic. The other 15 were given a similar quantity of distilled water. Upon completion of the the experiment, full bacterial counts were made from each vial. Summary statistics for the bacterial counts are provided below; assume that log-transformed counts are normally distributed.

Treatment	Mean of log counts	s.d. of log counts
Antibiotic	2.87	0.69
Distilled H20	3.27	0.86

(a) Provide a pooled estimate of s.d. (for observations on log scale)

```r
Sd.pool<-sqrt((((25-1)*(0.69^2))+(15-1)*(0.86^2)) /
                (25+15-2))
Sd.pool
```

```
## [1] 0.7570858
```

- **Answer:** The pooled standard deviation for this problem is 0.757

(b) Do these data suggest that the new antibiotic has an effect on bacterial counts in tissue cultures? (Answer this question by providing null and alternate hypothesis, provide the pooled t-ratio and p-value, and give an interprative conclusion - show all work).

Let μ_1 represent the log-transformed mean bacterial count of a culture given antibiotics, and let μ_2 represent the log-transformed mean bacterial count of a culture given distilled water.

- $H_0 : \mu_1 = \mu_2$
- $H_A : \mu 1 \neq \mu_2$

```r
# SE
SE<-Sd.pool * sqrt((1/25)+(1/15))

# T statistic
Mu.diff<-2.87-3.27
T.stat <- Mu.diff/SE # with 38 degrees of freedom
T.stat
```

```
## [1] -1.617709
```

```r
# p-value
# p-value
p<-2*pt(-abs(T.stat),df=38)
p
```

```
## [1] 0.1139974
```

- **Answer:** There is no evidence against the null hypothesis, and we conclude that the data do not provide any evidnce that the mean bacterial count was different between the bacteria treated with an antibiotic and distilled water.

(c) Provide a 90% confidence interval for the ratio between the median bacterial counts for the water and antibiotic treated cultures. Interpret the interval within the context given by this problem.

```r
CI.upper <- exp(((2.87-3.27) + (1.684*SE)))
CI.lower <- exp(((2.87-3.27) - (1.684*SE)))
CI.90<-data.frame(CI.lower,CI.upper)
CI.90
```

```
##    CI.lower CI.upper
## 1 0.4420239 1.016526
```

- **Answer:** We are 90% confident that the true difference in mean bacterial counts of cultures given antibiotics is 0.442 to 1.016 times that of cultures given distilled water.

3.) Dosing Guinea Pigs

The following boxplots and summary statistics are of the survival times *(in days)* of guinea pigs that were randomly assigned to a control group or a treatment group that received a dose of the tubercle bacilli. Scientists are interested to see if the treatment has an effect on the mean surival time.

(a) Write down the null and alternate hypotheses for testing the effect of treatment on the survival times.

- **Answer:** Let μ_1 represent the mean survival time (days) of mice given a dose of tubercle bacilli, and let μ_2 represent the suvival time (days) of mice that were not given a dose of tubercle bacilli.
- $H_0 : \mu_1 = \mu_2$
- $H_A : \mu 1 \neq \mu_2$

(b) Compute and report the Welch's t-ratio and degrees of freedom.

```r
SE <- sqrt((((117.931^2)/58)+((222.179^2)/64)))
T.ratio<- ((345.234-242.534)/SE)
T.ratio
```

```
## [1] 3.229794
```

```r
DF<- SE^4 /
          (((117.931^4)/((58^2)*(58-1))) +
             ((222.179^4)/((64^2)*(64-1))))
round(DF,1)
```

```
## [1] 97.8
```

- **Answer:** The Welch's t-ratio is 3.23 with 97.8 degrees of freedom.

(c) Use a t-table to provide a bound on the p-value (using nearest integer degrees of freedom).

- **Answer:** The p-value bounds are between 0.002 and 0.001

(d) Provide a scientific conclusion from the given context.

- **Answer:** There is strong evidence that the mean number of survival days of guinea pigs in the control group is greater than the mean number of survival days in the bacilli treatment group.

(e) Provide a 95% confidence interval for the difference in mean survival times and provide its interpretation.

```
CI.lower<- (345.234-242.534) - (1.984 * SE)
CI.upper<- (345.234-242.534) + (1.984 * SE)
CI<-data.frame(CI.lower,CI.upper)
CI
```

```
##   CI.lower CI.upper
## 1 39.61338 165.7866
```

- **Answer:** We are 95% confident that the mean survival time for guinea pigs in the control group is 39.61 to 63.09 days more than guinea pigs treated with tubercle bacilli.