Stat311 Homework 6 Template

Code

Joey Kang

2022-08-03

Read in the ice cream, birthweight, and cholesterol data sets.

Hide

```
IC.df <- read.csv("IceCream.csv", header=TRUE, as.is=TRUE)
IC.df$Sex <- as.factor(IC.df$Sex)
IC.df$Flavor <- as.factor(IC.df$Flavor)
#
BW.df <- read.csv("BirthWeight.csv", header=TRUE, as.is=TRUE)
BW.df$Smoker <- as.factor(BW.df$Smoker)
BW.df$BirthWt <- as.factor(BW.df$BirthWt)
#
C.df <- read.csv("Cholesterol.csv", header=TRUE, as.is=TRUE)
C.df$Cereal <- as.factor(C.df$Cereal)
```

Problem 1

Hide

```
mean_puzzle <- mean(IC.df$Puzzle)
sd_puzzle <- sd(IC.df$Puzzle)

z.test(IC.df$Puzzle,
  alternative = "two.sided",
  mu = mean_puzzle,
  sigma.x = sd_puzzle,
  conf.level = 0.95)$conf.int
```

```
## [1] 50.91712 53.89288
## attr(,"conf.level")
## [1] 0.95
```

The mean puzzle score for a population of 200 high school students falls between 50.91712 and 53.89288 with 95% confidence.

Problem 2

Hide

```
Chocolate <- filter(IC.df, Flavor == "2")
Strawberry <- filter(IC.df, Flavor == "3")

t.test(Chocolate$Video, Strawberry$Video, var.equal = FALSE, conf.level = 0.99)$conf.int
```

```
## [1] -12.976737  -2.515559
## attr(,"conf.level")
## [1] 0.99
```

We can say with 99% confidence that the mean difference in population mean video game scores for students that prefer chocolate and strawberry ice creams falls between -12.976737 and -2.515559.

Problem 3

Hide

```
# OMIT
```

Problem 4

Hide

```
binom.test(22, 42, conf.level = 0.95)$conf.int
```

```
## [1] 0.3641780 0.6799595
## attr(,"conf.level")
## [1] 0.95
```

We are 95% confident that the population difference between the proportion of mothers who smoked and mothers that did not smoke falls between about 36% and 68%.

Problem 5

Hide

```
prop.test(x=c(6,42), n=c(22,42),
          conf.level = 0.90, correct = FALSE)$conf.int
```

```
## [1] -0.8834540 -0.5710915
## attr(,"conf.level")
## [1] 0.9
```

We are 90% confident that the population difference between the proportion of mothers who smoked and mothers who did smoked who had low birthweight babies falls between about -0.8834540 and -0.5710915.

Problem 6

For the estimations problems 4 and 5, large sample conditions are met because for problem 4, both n p hat and n q hat are greater than 10. For problem 5, n1p1 and n1q1 are both greater than 10 and n2p2, n2q2 are both greater than 10.

Problem 7

Hide

```
# OMIT
```

Problem 8

Hide

```
tsum.test(mean.x = 22.5, s.x = 3.77, n.x = 25,
          mean.y = 25.2, s.y = 3.85, n.y = 18,
            alternative = "two.sided",
            mu = 0,
            conf.level = 0.95)$conf.int
```

```
## [1] -5.0920758 -0.3079242
## attr(,"conf.level")
## [1] 0.95
```

We are 95% confident that the population difference in mean Hamilton depression scale scores for treatment and placebo groups falls between about -5.0920758 and -0.3079242.