Homework 7

Kevin Jin

4/13/2022

Problem 1

```
# Calculate sample mean and standard deviation of the sample mean
x <- 180
n <- 15; total <- 3500
sigma <- 25
# P(X >= x)
pnorm(total, mean = x * n, sd = sqrt(n) * sigma, lower.tail = FALSE)
```

```
## [1] 7.140742e-17
```

The probability of a sample of 15 people exceeding 3,500 pounds total is close to zero at $7.1 * 10^{-17}$; it would be extremely unlikely and near impossible.

Problem 2

```
# Calculate parameters
n <- 30; total <- 600
lambda <- 25 # mean and variance
sd <- sqrt(lambda)
# P(X >= x)
pnorm(total, mean = lambda * n, sd = sqrt(n) * sd, lower.tail = FALSE)
```

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

The probability of selling more than 600 bottles in the next 30 days is 1 or very close to 1; it is essentially certain.

Problem 3

```
prop.test(x = 2, n = 30)
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  2 out of 30, null probability 0.5
## X-squared = 20.833, df = 1, p-value = 5.01e-06
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.01163184 0.23507287
## sample estimates:
```

```
##          p
## 0.06666667
```

We are 95% confident that the true proportion of left-handed students is between 1.16% and 23.5%; therefore, the 95% confidence interval does contain the actual value of 13.1%.

Problem 4

```
library(UsingR)
```

```
## Loading required package: MASS
## Loading required package: HistData
## Loading required package: Hmisc
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, units
##
## Attaching package: 'UsingR'
## The following object is masked from 'package:survival':
##
##     cancer
```

```
data(babies)
t.test(babies$age, babies$dage, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  babies$age and babies$dage
## t = -17.392, df = 1235, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.745356 -2.986035
## sample estimates:
## mean of the differences
##               -3.365696
```

The 95% confidence interval does not contain 0; therefore, we conclude that 0 is not a reasonable value for the true difference in mean mother's age and mean father's age.

For the paired t-test to work, we assume that the response variables (in this case, the age) are continuous, that both sets of observations are independent of each other, that both sets of observations follow a normal distribution, and that they do not contain outliers.