Part I

Exercise 1

a) The actual area between these two values is 1 - 2 * 0.1587 = 0.6826. The empirical rule says that this area should be 0.68.
b) The actual area between two standard deviations of the mean is 1 - 2 * 0.0228 = 0.9544 and the empirical rule says that this area should be 0.95. The actual area between three standard deviations of the mean is 1 - 2 * 0.0013 = 0.9974 and the empirical rule says that this area should be 0.997.
c) The 25th percentile is approximately 0.671 standard deviations below the mean. The 50th percentile is 0 standard deviations below the mean. The 75th percentile is approximately 0.671 standard deviations above the mean.

Exercise 2

a) Input: pnorm(65, mean = 69, sd = 2.8)
Output: 0.07656373, or 7.66%
b) Input: 1 - pnorm(75, mean = 69, sd = 2.8)
Output: 0.01606229, or 1.61%
c) Input: pnorm(72, mean = 69, sd = 2.8) - pnorm(66, mean = 69, sd = 2.8)
Output: 0.7160232 or 71.60%

Exercise 3

a) Input: qnorm(0.005, mean = 69, sd = 2.8)
Output: 61.78768 (inches)
b) Input: qnorm(0.9975, mean = 69, sd = 2.8)
Output: 76.85969 (inches)

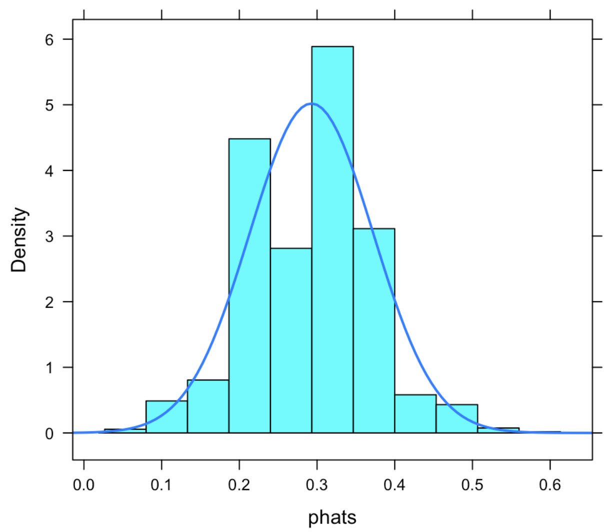
Exercise 4

a) Input:

```
pawnee <- read.csv("pawnee.csv")
n <- 30
N <- 541
M <- 1000
phats <- numeric(M)
set.seed(123)
for (i in seq_len(M)) {
  index <- sample(N, size = n)
  sample_i <- pawnee[index, ]
  phats[i] <- mean(sample_i$New_hlth_issue == "Y")
}

library(mosaic)
histogram(phats, fit = "normal")
```

Output:



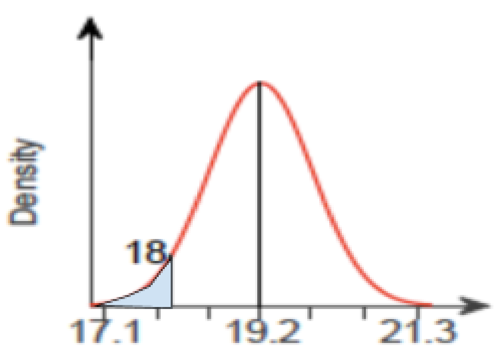
b) Input: mean(phats)
sd(phats)
Output: 0.2928 (mean)
0.07951963 (standard deviation)

c) The simulated distribution of sample proportions is approximately normal. It is unimodal and largely symmetric.

d) Input: mean(pawnee$New_hlth_issue == "Y")
Output: 0.2920518 (mean, theoretical)
Theoretical SD: sqrt((0.2921)*(1-0.2921)/30) = 0.083
Our empirical and theoretical means (0.2928 and 0.2921 respectively) and standard deviations (0.0795 and 0.083 respectively) are almost exactly the same.

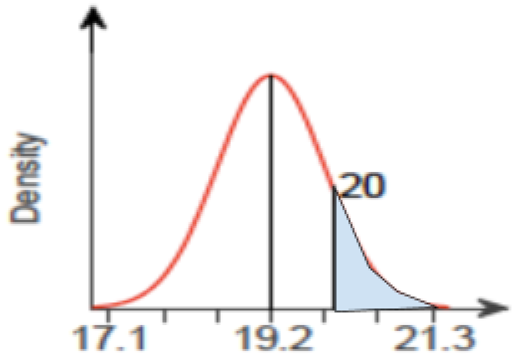
Part II

Exercise 1

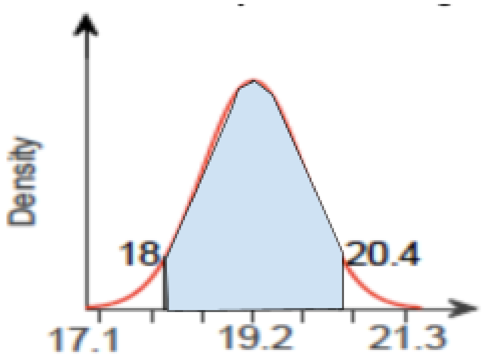


a)

The z-score is (18-19.2)/0.7 = -1.71. This would mean that 4.363% of newborns are shorter than 18 inches.

b) 

The z-score is (20-19.2)/0.7 = 1.14. This would mean that 1- 0.87286 or 12.714% of newborns are longer than 20 inches.

c) 

From part a) we know that 4.363% of newborns are shorter than 18 inches. The z-score corresponding to a raw value of 20.4 would be (20.4-19.2) = 1.71. We note that this is the absolute value of the z-score that corresponds to a raw value of 18. Therefore, 1 - 0.04363 * 2 = 91.274% of newborn babies are between 18 and 20.4 inches.

Exercise 2

The threshold for the top 30% would be the 70th percentile. The z-score that corresponds to a raw value of 428 would be (428-400)/60 = 0.47. Unfortunately, that z-score corresponds to the 68th percentile meaning that student won't be admitted.

Exercise 3

a) We should expect our sample proportion to be 58% or 0.58.
b) SE = sqrt(0.58*(1-0.58)/100) = 0.0493
c) We expect 58% of the students in the sample to have their driver's license, give or take 4.93%.
d) This would reduce the standard error. The new SE would be sqrt(0.58*(1-0.58)/100) = 0.0187.

Exercise 4

a) We expect 58% of the sample to watch television primarily through streaming services.
b) Condition 1: Random and independent observations are guaranteed in the problem statement.

Condition 2: np = 300*0.58 = 174 > 10 and n(1-p) = 300*0.42 = 126 > 10. The sample is sufficiently large.

Condition 3: 10n = 3000 and we can assume that there are way more than 3000 people in America who are between the ages of 18 and 29. The population is sufficiently large.

c) We have a normal distribution denoted by N(0.58, sqrt(0.58*0.42/300)) = N(0.58, 0.028). 181/300 = 0.603, which is less than one standard deviation from the mean. Therefore, finding 181 people in the sample who watch television primarily through streaming services wouldn't be surprising.

d) A raw value of 0.65 corresponds to a z-score = (0.65-0.58)/0.028 = 2.5 which corresponds to the 99.4th percentile. The probability that more than 65% of the sample watched television primarily through streaming services is 1 - 99.4 or 0.6%.

Exercise 5

a) Condition 1: Random and independent observations are guaranteed in the problem statement.

   Condition 2: np = 800*0.82 = 656 > 10, n(1-p) = 800*0.18 = 144 > 10. The sample is sufficiently large.

   Condition 3: 10n = 8000 and we can assume that there are more than 8000 adults in that country so the population is sufficiently large.

b) SE = sqrt(0.82*(1-0.82)/800) = 0.014. The z-score that corresponds to the 95th percentile is 1.96, so the margin of error = 0.014*1.96 = 0.02744. Therefore, our confidence interval is (0.82 - 0.02744, 0.82 + 0.02744) = (0.79256, 0.84744). We are 95% confident that the actual proportion of the adults in the country who believe that protecting the rights of those with unpopular views is a very important component of a strong democracy lies within that interval.

c) The confidence interval would be narrower. The z-score that corresponds to the 90th percentile is lower than the z-score that corresponds to the 95th percentile and since ME = Z * SE, a smaller z-score means a smaller margin of error means a narrower interval.