It is given that $EXERCISE 4$

If you use $z$ or if you use $\hat{p}$.

Both $np$ and $n(1-p)$ are $\geq 5$. Also, it is reasonable to assume that these $400$ voters is less than $10\%$ of the entire population.

EXERCISE 2

It is given that $p = 0.92$, $n = 160$. We want $P(\hat{p} > 0.95)$. This is equal to:

$$P(\hat{p} > 0.95) = P(Z > \frac{0.95 - 0.92}{\sqrt{\frac{0.92(1-0.92)}{160}}} = P(Z > 1.40) = 1 - 0.9192 = 0.0808.$$  

The assumptions hold: Large sample, both $np$ and $n(1-p)$ are $\geq 10$. Also, it is reasonable to assume that these $160$ is less than $10\%$ of the entire population. However, the independence assumption may not hold.

EXERCISE 3

It is given $p = 0.60$, $n = 120$. Very sure means that the probability that the restaurant does not have enough seats for nonsmokers is very small - say $5\%$. Therefore, we need to find the sample proportion above which the probability is only $5\%$. This corresponds to a $z = 1.645$.

$$1.645 = \frac{\hat{p} - 0.60}{\sqrt{\frac{0.60(1-0.60)}{120}}} \Rightarrow \hat{p} = 0.674.$$  

Therefore, they should have $0.674(120) \approx 81$ seats in the smoke-free area. The assumptions should hold: Both $np$, and $n(1-p)$ are $\geq 10$, customers must be independent of each other.

Note: If you use other $z$ you will get a different answer. For example, if you use $z = 2$ you will get $84$ seats, or if you use $z = 3$ you will get $90$ seats.

EXERCISE 4

It is given that $\mu = 2.9gm/mi$, $\sigma = 0.4gm/mi$, and $n = 80$.

a. The distribution of the sample mean $\bar{X}$ according to the central limit theorem is:

$$\bar{X} \sim N(2.9, \frac{0.4}{\sqrt{80}}) \text{ or } \bar{X} \sim N(2.9, 0.045).$$  

b. $P(3.0 < \bar{X} < 3.1) = P(\frac{3.0 - 2.9}{0.045} < Z < \frac{3.1 - 2.9}{0.045}) = P(2.22 < Z < 4.44) \approx 1 - 0.9868 = 0.0132.$$

c. We want to find the 5th percentile of the distribution of $\bar{X}$. The corresponding $z$ value is -1.645. Therefore, $-1.645 = \frac{\bar{X} - 2.9}{0.045} \Rightarrow \bar{X} = 2.83gm/mi.$

EXERCISE 5

It is given that the amount of water consumed follows $N(19, 5)$. There are $n = 100$ employees.

a. The total amount of water demanded by the employees follows $T \sim N(100(19), 5\sqrt{100})$ or $T \sim N(1900, 50)$. We want $P(T > 2000) = P(Z > \frac{2000 - 1900}{50}) = P(Z > 2) = 1 - 0.9772 = 0.0228.$

b. This is binomial with $n = 4, p = 0.0228$.

We want $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{4}{0}(0.0228)^0(1 - 0.0228)^4 = 0.0881.$
c. Here, we have a binomial problem with \( n = 365, p = 0.0228 \). We want \( P(X > 15) \). We will use the normal approximation to binomial, with \( \mu = 365(0.0228) = 8.32, \sigma = \sqrt{365(0.0228)(1-0.0228)} = 2.85 \). And the desired probability is \( P(X > 15) = P(Z > \frac{15.5-8.32}{2.85}) = P(Z > 2.52) = 1 - 0.9941 = 0.0059 \).

**EXERCISE 6**

Let \( X \) be the casino’s net payoff. Then the probability distribution of \( X \) is

\[
\begin{array}{c|c}
 X & P(X) \\
\hline
-35 & \frac{38}{94} \\
1 & \frac{38}{94}
\end{array}
\]

This distribution has \( \mu = 0.05263 \) and \( \sigma = 5.76 \).

a. \( P(T > 400) = P(Z > \frac{400 - 10000(0.05263)}{5.76\sqrt{10000}}) = P(Z > -0.22) = 0.5871 \).

b. \( P(T > 400) = P(Z > \frac{400 - 9000(0.05263)}{5.76\sqrt{9000}}) = P(Z > -2.51) = 0.9940 \).

c. Let \( k \) be the minimum number of games the casino must win. To find \( k \) solve: \(-35(10000 - k) + 1k > 400 \Rightarrow k \geq 9733.3 \approx 9734 \).

\[
P(Y \geq 9734) = \sum_{y=9734}^{10000} \binom{10000}{y} \frac{37}{38}^y \frac{1}{38}^{10000-y}.
\]

**EXERCISE 7**

It is given that \( X \sim N(180, 32) \). A sample of \( n = 64 \) is selected. Then \( \bar{X} \sim N(180, \frac{32}{\sqrt{64}}) \) or \( \bar{X} \sim N(180, 4) \).

a. \( P(\bar{X} > 185) = P(Z > \frac{185 - 180}{4}) = P(Z > 1.25) = 1 - 0.8944 = 0.1056 \).

b. \( P(\bar{X} < 173) = P(Z < \frac{173 - 180}{4}) = P(Z < -1.75) = 0.0401 \).

c. \( P(173 < \bar{X} < 185) = P(-1.75 < Z < 1.25) = 0.8944 - 0.0401 = 0.8543 \).

**EXERCISE 8**

The model is \( y_i = \beta_1 x_i + \epsilon_i \). Because \( E(\epsilon_i) = 0 \) and \( var(\epsilon_i) = \sigma^2 \) it follows that \( E(y_i) = \beta_1 x_i \) and \( var(y_i) = \sigma^2 \). We can write \( \hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \) as \( \hat{\beta}_1 = x_1 \hat{\beta}_1 x_1 + x_2 \hat{\beta}_2 x_2 + \ldots + x_n \hat{\beta}_n x_n \).

a. \[
E(\hat{\beta}_1) = E\left( \frac{x_1 y_1}{\sum_{i=1}^{n} x_i^2} + \frac{x_2 y_2}{\sum_{i=1}^{n} x_i^2} + \ldots + \frac{x_n y_n}{\sum_{i=1}^{n} x_i^2} \right) = \frac{x_1 E(y_1)}{\sum_{i=1}^{n} x_i^2} + \frac{x_2 E(y_2)}{\sum_{i=1}^{n} x_i^2} + \ldots + \frac{x_n E(y_n)}{\sum_{i=1}^{n} x_i^2} = \beta_1 \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} = \beta_1.
\]

b. \[
\text{var}(\hat{\beta}_1) = \text{var}\left( \frac{x_1 y_1}{\sum_{i=1}^{n} x_i^2} + \frac{x_2 y_2}{\sum_{i=1}^{n} x_i^2} + \ldots + \frac{x_n y_n}{\sum_{i=1}^{n} x_i^2} \right) = \frac{x_1^2 \text{var}(y_1)}{(\sum_{i=1}^{n} x_i^2)^2} + \frac{x_2^2 \text{var}(y_2)}{(\sum_{i=1}^{n} x_i^2)^2} + \ldots + \frac{x_n^2 \text{var}(y_n)}{(\sum_{i=1}^{n} x_i^2)^2} = \sigma^2 \frac{\sum_{i=1}^{n} x_i^2}{(\sum_{i=1}^{n} x_i^2)^2} = \sigma^2 \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2}.
\]

Finally, because \( \epsilon_i \sim N(0, \sigma) \) we get the distribution \( \hat{\beta}_1 \sim N(\beta_1, \frac{\sigma}{\sqrt{\sum_{i=1}^{n} x_i^2}}) \).