Exercise #8

Question 1:

The National Health and Nutrition Examination Survey of 1988-1994(NHANES III, A-1) estimated the mean serum cholesterol level for U.S. females aged 20-74 years to be 204 mg/dl. The estimate of the standard deviation was approximately 44. Using these estimates as the mean μ and the standard deviation σ for the U.S. population, consider the sampling distribution of the sample mean based on samples of size **50** drawn from women of this age group. Find:

Given n=50(large) μ=204 $\sigma = 44$

Then, $\bar{X} \approx Normal(\mu, \frac{\sigma^2}{n})$

1) **The mean of the sampling distribution $\bar{X}$**

   $\text{Mean}(\bar{X}) = \mu_{\bar{X}} = \mu = 204$

2) **The standard error of $\bar{X}$**

   $$S.E = \frac{\sigma}{\sqrt{n}} = \frac{44}{\sqrt{50}} = 6.22$$

3) **P($\bar{X}$< 200)**

$$= P\left(Z < \frac{200 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z < \frac{200 - 204}{\frac{44}{\sqrt{50}}}\right)$$

$$= P(Z < -0.64)$$

$$= 0.26109$$

4) **P(170 < $\bar{X}$< 195)**

$$= P\left(\frac{170-\mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{195-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(\frac{170 - 204}{\frac{44}{\sqrt{50}}} < Z < \frac{195 - 204}{\frac{44}{\sqrt{50}}}\right)$$

$$= P(-5.46 < Z < -1.45)$$

$$= P(Z < -1.45) - P(Z < -5.46)$$

$$= 0.07353 - 0 = 0.07353$$

5) $P(\bar{X} < 207)$

$$= P\left(Z < \frac{207 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z < \frac{207 - 204}{\frac{44}{\sqrt{50}}}\right)$$

$$= P(Z < 0.48)$$

$$= 0.68439$$

6) $P(\bar{X} > 190)$

$$= P\left(Z > \frac{190 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z > \frac{190 - 204}{\frac{44}{\sqrt{50}}}\right)$$

$$= P(Z > -2.25)$$

$$= 1 - P(Z < -2.25)$$

$$= 1 - 0.01222 = 0.98778$$

Question 2:

If the uric acid values in normal adult males are approximately normally distributed with a mean and standard deviation of 5.7 and 1mg, respectively, for a sample of size 9 find:

Given, normally distributed n=9 μ=5.7 $\sigma = 1$

$$\bar{X} \sim Normal\left(\mu, \frac{\sigma^2}{n}\right)$$

1) **The mean of $\bar{X}$**

Mean($\bar{X}$) $= \mu_{\bar{X}} = \mu = 5.7$

2) **The standard error of $\bar{X}$**

$$S.E = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{9}} = \frac{1}{3} = 0.33$$

3) The probability that **the mean** of the uric acid values is greater than 6

$$P(\bar{X} > 6) = P\left(Z > \frac{6-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(Z > \frac{6-5.7}{\frac{1}{\sqrt{9}}}\right)$$

$$= P(Z > 0.9)$$

$$= 1 - P(Z < 0.9)$$

$$= 1 - 0.81594 = 0.18406$$

4) The probability that **the mean** of the uric acid values is between 5 and 6.

$$P(5 < \bar{X} < 6) = P\left(\frac{5-\mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{6-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(\frac{5-5.7}{\frac{1}{\sqrt{9}}} < Z < \frac{6-5.7}{\frac{1}{\sqrt{9}}}\right)$$

$$= P(-2.1 < Z < 0.9)$$

$$= P(Z < 0.9) - P(Z < -2.1)$$

$$= 0.81594 - 0.01786 = 0.79808$$

5) The probability that **the mean** of the uric acid values is less than 5.8

$$P(\bar{X} < 5.8) = P\left(Z < \frac{5.8-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z < \frac{5.8-5.7}{\frac{1}{\sqrt{9}}}\right)$$

$$= P(Z < 0.30) = 0.61791$$

Question 3(H.W):

Given a normally distributed population with mean of 100 and a standard deviation of 20, find the following based on a sample of size 16:

1) The mean of $\overline{X}$ (answer =100)

2) The variance of $\overline{X}$ (Answer = 25)

3) The standard error of $\overline{X}$ (Answer = 5)

4) $P(\bar{X} \geq 100) =$ (Answer = 0.5)

5) $P(\bar{X} \leq 110)$ (Answer = 0.97725)

6) $P(96 < \bar{X} < 108)$ (Answer = 0.73334)

H.W

Given μ = 50, σ = 16, and n = 64, find:

1) $P(45 < \bar{x} < 55)$ (Answer = 0.98758)
2) $P(\bar{x} > 53)$ (Answer = 0.06681)
3) $P(\bar{x} < 47)$ (Answer = 0.06681)
4) $P(49 < \bar{x} < 56)$ (Answer = 0.69011)

Question 4:

In a study, the data about the serum cholesterol level in U.S. females are given in the following table:

Population	Age	Mean	Standard Deviation
A	30 – 59	189	34.7
B	20 – 29	183	37.2

Suppose we select a simple random sample **of size 50** independently from each population, then:

1) The mean of $\bar{x}_A - \bar{x}_B$ is:

   $\text{Mean}(\bar{x}_A - \bar{x}_B) = \mu_A - \mu_B = 189 - 183 = 6$

2) The standard error of $\bar{x}_A - \bar{x}_B$ is:

$$\text{S.E} = \sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}} = \sqrt{\frac{34.7^2}{50} + \frac{37.2^2}{50}} = 7.194$$

3) The distribution of $\bar{x}_A - \bar{x}_B$ is:

$$\bar{x}_A - \bar{x}_B \sim Normal(\mu_A - \mu_B\,, \frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B})$$

$$\bar{x}_A - \bar{x}_B \sim Normal(6\,, 51.76)$$

4) The probability that the difference between sample means $\bar{x}_A - \bar{x}_B$ will be mor than 8

$$P(\bar{x}_A - \bar{x}_B > 8) = P\left(Z > \frac{8-(\mu_A-\mu_B)}{\sqrt{\frac{\sigma_A{}^2}{n_A}+\frac{\sigma_B{}^2}{n_B}}}\right)$$

$$= P\left(Z > \frac{8-(189-183)}{\sqrt{\frac{34.7^2}{50}+\frac{37.2^2}{50}}}\right)$$

$$= \quad P(Z>0.28) = P(Z < -0.28) = 0.38974$$

Question 5:

In a study, the calcium levels in men and women ages 60 years or older are summarized in the following table:

		Mean	Standard Deviation
1	**Men**	797	482
2	**Women**	660	414

If we take a random sample of 40 men and 35 women, then:

1) The mean of $\bar{x}_1 - \bar{x}_2$ is :

Mean $(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$ =797 - 660 =137

2) The variance of $\bar{x}_1 - \bar{x}_2$ is:

$$\text{Variance}(\bar{x}_1 - \bar{x}_2) = \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2} = \frac{482^2}{40} + \frac{414^2}{35} = 10705.13$$

3) The distribution of $\bar{x}_1 - \bar{x}_2$ is:

$$\bar{x}_1 - \bar{x}_2 \sim Normal(\mu_1 - \mu_2, \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2})$$

$$\bar{x}_1 - \bar{x}_2 \sim Normal(137, 10705.13)$$

4) The probability of obtaining a difference between sample means $\bar{x}_1 - \bar{x}_2$ of 100mg or less.

$$P(\bar{x}_1 - \bar{x}_2 < 100) = P(Z < \frac{100 - (\mu_1 - \mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}})$$

$$= P(Z < \frac{100 - (797 - 660)}{\sqrt{\frac{482^2}{40} + \frac{414^2}{35}}})$$

$$= P(Z < -0.36) = 0.35942$$

Question 6:

Smith et al. (A-5) performed a retrospective analysis of data on 782 eligible patients admitted with myocardial infarction to a 46-bed cardiac service facility. Of these patients, 248 reported a past myocardial infarction. Suppose 50 subjects are chosen at random from the population, what is:

Given that n=50 $p = \frac{248}{782} = 0.317$ then $q = 1 - p$ =0.683

1) The mean of $\hat{P}$

Mean of $\hat{P} = P = \frac{248}{782} = 0.317$

2) The variance of $\hat{P}$

Variance of $\hat{P} = \frac{pq}{n} = \frac{0.317\times0.683}{50} = 4.3\times 10^{-3} = 0.0043$

3) The standard error of $\hat{P}$

Standard error of $\hat{P} = \sqrt{\frac{pq}{n}} = 0.0658$

4) The distribution of $\hat{P}$

$\hat{P} \sim Normal(\text{ p}, \frac{pq}{n})$

$\hat{P} \sim Normal(0.317, 0.0043)$

5) The probability that over **40 percent** would report previous myocardial infarctions?

$P(\hat{P} > 0.4) = P(Z > \frac{0.4-p}{\sqrt{\frac{pq}{n}}})$

$= P(Z > \frac{0.4-0.317}{\sqrt{\frac{0.317\times0.683}{50}}}) = P(Z > 1.26) = 0.10383$

Question 7:

Researchers estimated that 64 percent of U.S. adults ages 20-74 were overweight or obese. Use this estimate as the population proportion for U.S. adults ages 20-74. If 125 subjects are selected at random from the population, what is:

given that n=125 and $p = \frac{64}{100} = 0.64$ then $q = 0.36$

1) The mean of $\hat{P}$

Mean of $\hat{P} = p = \frac{64}{100} = 0.64$

2) The variance of $\hat{P}$

variance of $\hat{P} = \frac{pq}{n} = \frac{(0.64)(0.36)}{125} = 0.00184$

3) The standard error of $\hat{P}$

standard error of $\hat{P} = \sqrt{\frac{pq}{n}}$ =0.0429

4) The distribution of $\hat{P}$

$\hat{P} \sim Normal(\text{p}, \frac{pq}{n})$

$\hat{P}$ ~ Normal(0.64 , 0.00184)

5) The probability that **70 percent** or more would be found to be overweight or obese?

$$P(\hat{P} > 0.7) = P(Z > \frac{0.7 - p}{\sqrt{\frac{pq}{n}}})$$

$$= P(Z > \frac{0.7 - 0.64}{\sqrt{\frac{(0.64)(0.36)}{125}}}) \; P(Z > 1.397) = P(Z > 1.40)$$

$$= 1 - P(Z < 1.40) = 1 - 0.91924 = 0.08076$$

H.W

given a population in which p=0.6 and a random sample from this population of size 100 find:

1) P ($\hat{P} \geq 0.65$) (Answer = 0.15386)
2) P ($0.56 \leq \hat{P} \leq 0.63$) (Answer = 0.52296)
3) P ($\hat{P} \leq 0.58$) (Answer = 0.3409)

H.W

It is known that 35 percent of the members of a certain population suffer from one or more chronic diseases. What is the probability that in a sample of 200 subjects drawn at random from this population 40 percent or more will have at least one chronic disease?

(Answer = 0.0694)

$$P(\hat{P} \geq 0.4) = P\left(Z \geq \frac{0.4-0.35}{\sqrt{\frac{0.35x0.65}{200}}}\right) = P(Z > 1.48) =$$

$$1 - P(Z < 1.48) = 1 - 0.93056 = 0.06944$$

Question 8:

In a study, the Census Bureau stated that for Americans in the age group 18 to 24 years, 64.8 percent had private health insurance. In the age group 25-34years, the percentage was 72.1. Assume that these percentages are the population parameters in those age groups for the United States. Suppose we select a random sample of 250 Americans from the 18–24 age group and an independent random sample of 200 Americans from the age group 25–34; find:

Geven that $n_1 = 250$ and $p_1 = \frac{64.8}{100} = 0.648$ then $q_1 = 1 - 0.648 = 0.352$

And $n_2 = 200$ and $p_2 = \frac{72.1}{100} = 0.721$ then $q_1 = 1 - 0.721 = 0.279$

1) the mean of $\hat{P}_1 - \hat{P}_2$

mean of $\hat{p}_1 - \hat{p}_2 = p_1 - p_2 = 0.648 - 0.721 = -0.073$

2) the standard error of $\hat{P}_1 - \hat{P}_2$

$$S.E = \hat{P}_1 - \hat{P}_2 = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} = \sqrt{\frac{0.648 \times 0.352}{250} + \frac{0.721 \times 0.279}{200}} = 0.0438$$

3) The distribution of $\hat{P}_1 - \hat{P}_2$

Distribution of $\hat{P}_1 - \hat{P}_2$ ~Normal$(p_1 - p_2\ ,\ \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2})$

$\hat{P}_1 - \hat{P}_2$ ~Normal$(-0.073\ ,\ 1.9 \times 10^{-3})$

4) P $(\hat{P}_1 - \hat{P}_2 \leq 0.3)$

$$p(Z \leq \frac{0.3-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}) = p(Z \leq \frac{0.3-(-0.073)}{\sqrt{\frac{(0.648)(0.352)}{250}+\frac{(0.721)(0.279)}{200}}})$$

$$= P(Z \leq 8.52) = 1$$

5) The probability that $\hat{P}_1 - \hat{P}_2$ is less than 6 percent.

$$P(\hat{P}_1 - \hat{P}_2 < 0.06) = p(Z < \frac{0.06-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}) = p(Z < \frac{0.06-(-0.073)}{\sqrt{\frac{(0.648)(0.352)}{250}+\frac{(0.721)(0.279)}{200}}})$$

$$= P(Z \leq 3.04) = = 0.99882$$

H.W

From the results of a survey conducted by the U.S. Bureau of Labor Statistics (A-9), it was estimated that 21 percent of workers employed in the Northeast participated in health care benefits programs that included vision care. The percentage in the South was 13 percent. Assume these percentages are population parameters for the respective U.S. regions. Suppose we select a simple random sample of size 120 northeastern workers and an independent simple random sample of 130 southern workers. Find:

1) The mean of $\hat{P}_1 - \hat{P}_2$ (Answer = 0.08)
2) The standard error of $\hat{P}_1 - \hat{P}_2$ (Answer = 0.0475)
3) The distribution of $\hat{P}_1 - \hat{P}_2$ (Answer = N(0.08,0.0023))
4) What is the probability that the difference between sample proportions, $\hat{P}_1 - \hat{P}_2$, will be between 0.04 and 0.20.

(Answer = 0.79385)