APMA 1650 – Homework 8

1. Section 7.2: 7.9

   Refer to Example 7.2. The amount of fill dispensed by a bottling machine is normally distributed with $\sigma = 1$ ounce. If $n = 9$ bottles are randomly selected from the output of the machine, we found that the probability that the sample mean will be within 0.3 ounce of the true mean is .6318. Suppose that $\bar{Y}$ is to be computed using a sample of size n.

   (a) a
   If $n = 16$, what is $P(|\bar{Y} - \mu| \leq .3)$?
   See Example 7.2. By Theorem 7.1, $\bar{Y}$ has a normal sampling distribution with mean μ and variance $\sigma^2/n = 1/16$. $(\bar{Y} - \mu)/(\sigma/\sqrt{n})$ has a standard normal distribution, so

   $$P(|\bar{Y} - \mu| \leq .3) = P(-\frac{0.3}{1/\sqrt{16}} \leq Z \leq \frac{0.3}{1/\sqrt{16}}) = P(-1.2 \leq Z \leq 1.2) = .7698$$

   (b) b
   Find $P(|\bar{Y} - \mu| \leq .3)$ when $\bar{Y}$ is to be computed using samples of sizes $n = 25$, $n = 36$, $n = 49$, and $n = 64$.
   Similar to the previous problem, we find

   $$P(|\bar{Y} - \mu| \leq .3) = P(-.3\sqrt{n} \leq Z \leq .3\sqrt{n}) = 1 - 2P(Z > .3\sqrt{n})$$

   For $n = 25, 36, 69,$ and 64, the probabilities are (respectively) .8664, .9284, .9642, and .9836.

   (c) c
   What pattern do you observe among the values for $P(|\bar{Y} - \mu| \leq .3)$ that you observed for the various values of n?
   The probabilities increase with n, which is intuitive since the variance of $\bar{Y}$ decreases with n.

2. Section 7.2: 7.11

   A forester studying the effects of fertilization on certain pine forests in the Southeast is interested in estimating the average basal area of pine trees. In studying basal areas of similar trees for many years, he has discovered that these measurements (in square inches) are normally distributed with standard deviation approximately 4 square inches. If the forester samples $n = 9$ trees, find the probability that the sample mean will be within 2 square inches of the population mean.

Similar to the previous problem, the variance of $\bar{Y}$ is σ^2/n, so

$$\begin{aligned} P(|\bar{Y}-\mu| \leq 2) &= P(-\frac{2}{4/\sqrt{9}} \leq Z \leq \frac{2}{4/\sqrt{9}}) \\ &= P(-1.5 \leq Z \leq 1.5) \\ &= 1 - 2P(Z > 1.5) \\ &= 1 - 2(0.0668) \\ &= .8664 \end{aligned}$$

3. Section 7.2: 7.15

Suppose that $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ are independent random samples, with the variables X_i normally distributed with mean μ_1 and variance σ_1^2 and the variables Y_i normally distributed with mean μ_2 and variance σ_2^2. The difference between the sample means, $\overline{X} - \overline{Y}$, is then a linear combination of $m + n$ normally distributed random variables and, by Theorem 6.3, is itself normally distributed.

We use Theorems 6.3 and 7.1 for the following problems. In particular, we already know that $\bar{X}$ has mean μ_1 and variance σ_1^2/n, and $\bar{Y}$ has mean μ_2 and variance σ_2^2/n.

(a) Find $E(\overline{X} - \overline{Y})$.
$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2$.

(b) Find $V(\overline{X} - \overline{Y})$.
Because X_i and Y_j are independent for all i, j, $\bar{X}$ is a function of X_i, and $\bar{Y}$ is a function of Y_j, $\bar{X}$ is independent from $\bar{Y}$. We know that for two independent random variables U and V, $Var(U + V) = Var(U) + Var(V)$. Also, $Var(-U) = |-1|^2 Var(U) = Var(U)$. Therefore, $Var(\bar{X} - \bar{Y}) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$

(c) Suppose that $\sigma_1^2 = 2$, $\sigma_2^2 = 2.5$, and $m = n$. Find the sample sizes so that $(\overline{X} - \overline{Y})$ will be within 1 unit of $(\mu_1 - \mu_2)$ with probability .95.
It is required that $P(|\bar{X} - \bar{Y} - (\mu_1 - \mu_2)| \leq 1) = .95$. Using the result in part b for standardization with $n = m$, $\sigma_1^2 = 2$, and $\sigma_2^2 = 2.5$, we obtain

$$\begin{aligned} 0.95 &= P(|\bar{X} - \bar{Y} - (\mu_1 - \mu_2)| \leq 1) \\ &= P(-\frac{1}{\sqrt{2/n + 2.5/n}} \leq \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{2/n + 2.5/n}} \leq \frac{1}{\sqrt{2/n + 2.5/n}}) \\ &= P(-\frac{1}{\sqrt{4.5/n}} \leq Z \leq \frac{1}{\sqrt{4.5/n}}) \end{aligned}$$

We know that the probability Z lies between -2 and 2, or is at most 2 standard deviations from the mean, is about 95%. This is achieved when $n = 18$.

4. Section 7.2: 7.21

   Refer to Exercise 7.13. Suppose that $n = 20$ observations are to be taken on ln(LC50) measurements and that $\sigma^2 = 1.4$. Let S^2 denote the sample variance of the 20 measurements.

   The values below can be found by using percentiles from the chi-square distribution. With $\sigma^2 = 1.4$ and $n = 20$, $\frac{19S^2}{1.4}$ has a chi-square distribution with 19 degrees of freedom.

   (a) Find a number b such that $P(S^2 \le b) = .975$.
   $P(S^2 \le b) = P\left(\frac{(n-1)S^2}{\sigma^2} \le \frac{(n-1)b}{\sigma^2}\right) = P\left(\frac{19S^2}{1.4} \le \frac{19b}{1.4}\right) = .975$. It must be true that $\frac{19b}{1.4} = 32.8523$, the 97.5%-tile of this chi-square distribution, and so $b = 2.42$.

   (b) Find a number a such that $P(a \le S^2) = .975$.
   Similarly, $P(S^2 \ge a) = P\left(\frac{(n-1)S^2}{\sigma^2} \ge \frac{(n-1)a}{\sigma^2}\right) = .974$. Thus, $\frac{19a}{1.4} = 8.96055$, the 2.5%-tile of this chi-square distribution, and so $a = .656$.

   (c) If a and b are as in parts (a) and (b), what is $P(a \le S^2 \le b)$?
   $P(a \le S^2 \le b) = .95$.

5. Section 7.2: 7.30a

   Suppose that Z has a standard normal distribution and that Y is an independent χ^2-distributed random variable with ν degrees of freedom (df). Then, according to Definition 7.2,

   $$T = \frac{Z}{\sqrt{Y/\nu}}$$

   has a t distribution with ν df.

   (a) If Z has a standard normal distribution, give $E(Z)$ and $E(Z^2)$. *Hint: For any random variable,* $E(Z^2) = Var(Z) + (E(Z))^2$.
   $E(Z) = 0$, $E(Z^2) = V(Z) + [E(Z)]^2 = 1$

6. Section 7.3: 7.42

   The fracture strength of tempered glass averages 14 (measured in thousands of pounds per square inch) and has standard deviation 2.

   Let $\bar{Y}$ denote the sample mean strength of 100 randomly selected pieces of glass. Then the quantity $(\bar{Y} - 14)/(2/\sqrt{100})$ has an approximate standard normal distribution.

   (a) What is the probability that the average fracture strength of 100 randomly selected pieces of this glass exceeds 14.5?
   $P(\bar{Y} > 14.5) = P(Z > (14.5 - 14)/(2/\sqrt{100})) = P(Z > 2.5) = .0062$

(b) Find an interval that includes, with probability 0.95, the average fracture strength of 100 randomly selected pieces of this glass.

We have that $P(-1.96 < Z < 1.96) = .95$. So $-1.96 = (a-14)/.2 \implies a = 13.608$ and $1.96 = (b-14)/.2 \implies b = 14.392$ So the interval is [13.608,14.392]

7. Section 7.3: 7.43

An anthropologist wishes to estimate the average height of men for a certain race of people. If the population standard deviation is assumed to be 2.5 inches and if she randomly samples 100 men, find the probability that the difference between the sample mean and the true population mean will not exceed .5 inch.

Let $\bar{Y}$ denote the mean height and $\sigma = 2.5$ inches. By CLT

$$P(|\bar{Y}-\mu| \leq 0.5) = P(-0.5 \leq \bar{Y}-\mu \leq 0.5) \approx P\left(\frac{-0.5(10)}{2.5} \leq Z \leq \frac{0.5(10)}{2.5}\right)$$

$$= P(-2 \leq Z \leq 2) = .9544$$

8. Section 7.3: 7.45

Workers employed in a large service industry have an average wage of \$ 7.00 per hour with a standard deviation of \$0.50. The industry has 64 workers of a certain ethnic group. These workers have an average wage of \$6.90 per hour. Is it reasonable to assume that the wage rate of the ethnic group is equivalent to that of a random sample of workers from those employed in the service industry? [Hint: Calculate the probability of obtaining a sample mean less than or equal to \$6.90 per hour.]

Let $\bar{Y}$ denote the mean wage calculated from a sample of 64 workers. Then,

$$P(\bar{Y} \leq 6.90) \approx P(Z \leq \frac{\sqrt{64}(6.9-7)}{.5}) = P(Z \leq -1.60) = .0548$$

9. Section 7.3: 7.58

Suppose that $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_n$ are independent random samples from populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. Show that the random variable

$$U_n = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{(\sigma_1^2+\sigma_2^2)/n}}$$

satisfies the conditions of Theorem 7.4 and thus that the distribution function of U_n converges to a standard normal distribution function as $n \to \infty$. [Hint: Consider $W_i = X_i - Y_i$, for $i = 1, 2, ..., n$.]

$W_i = X_i - Y_i \implies E(W_i) = E(X_i) - E(Y_i) = \mu_1 - \mu_2$ while for the variance $V(W_i) = V(X_i) - V(Y_i) = \sigma_1^2 - \sigma_2^2$ since X and Y are independent. Thus

$$\bar{W} = \frac{1}{n}\sum_{i=1}^{n} W_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \bar{X} - \bar{Y}$$

Therefore $E(\bar{W}_i) = \mu_1 - \mu_2$ and $V(\bar{W}_i) = \frac{1}{n}(\sigma_1^2 - \sigma_2^2)$. Now since the W_i are independent,

$$U_n = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} = \frac{\bar{W} - E(\bar{W})}{\sqrt{V(\bar{W})}}$$

Which satisfies the conditions of Theoreom 7.4 and has a limiting standard normal distribution