Cavallaro, Jeffery
Math 161A
Homework #7

5.1

A service station has both self-service and full-service islands. On each island, there is a singular regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, and let Y denote the number of hoses on the full-service island in use at that time. The joint pmf of X and Y appears in the accompanying tabulation.

x \ y	0	1	2
0	0.10	0.04	0.02
1	0.08	0.20	0.06
2	0.06	0.14	0.30

a) What is $P(X = 1 \text{ and } Y = 1)$?

$$P(X = 1 \text{ and } Y = 1) = 0.20$$

b) Compute $P(X \leq 1 \text{ and } Y \leq 1)$?

$$P(X \leq 1 \text{ and } Y \leq 1) = 0.10 + 0.04 + 0.08 + 0.20 = 0.42$$

c) Give a word description of the event $\{X \neq 0 \text{ and } Y \neq 0\}$, and compute the probability of this event.

The event that at least one hose is in use on each island.

$$P(X \neq 0 \text{ and } Y \neq 0) = 0.20 + 0.06 + 0.14 + 0.30 = 0.70$$

d) Compute the marginal pmf of X and of Y. Using $p_X(x)$, what is $P(X \leq 1)$?

x \ y	0	1	2	$p_X(x)$
0	0.10	0.04	0.02	0.16
1	0.08	0.20	0.06	0.34
2	0.06	0.14	0.30	0.50
$p_Y(y)$	0.24	0.38	0.38	

$$P(X \leq 1) = 0.16 + 0.34 = 0.50$$

e) Are X and Y independent rv's? Explain.

No.

$$p_X(0)p_Y(0) = 0.16 \cdot 0.24 = 0.384 \neq 0.10 = p(0,0)$$

5.2

A large but sparsely populated county has two small hospitals, one at the south end of the county and the other at the north end. The south hospital's emergency room has four beds, whereas the north hospital's emergency room has only three beds. Let X denote the number of south beds occupied at a particular time on a given day, and let Y denote the number of north beds occupied at the same time on the same day. Suppose that these two rv's are independent; that the pmf of X puts probability masses 0.1, 0.2, 0.3, 0.2, and 0.2 on the x values 0, 1, 2, 3, and 4, respectively; and that the pmf of Y distributes probabilities 0.1, 0.3, 0.4, and 0.2 on the y values 0, 1, 2, and 3, respectively.

a) Display the joint pmf of X and Y in a joint probability table.

y \ x	0	1	2	3	4	$p_Y(y)$
0	0.01	0.02	0.03	0.02	0.02	0.10
1	0.03	0.06	0.09	0.06	0.06	0.30
2	0.04	0.08	0.12	0.08	0.08	0.40
3	0.02	0.04	0.06	0.04	0.04	0.20
$p_X(x)$	0.10	0.20	0.30	0.20	0.20	

b) Compute $P(X \leq 1 \text{ and } Y \leq 1)$ by adding probabilities from the joint pmf, and verify that this equals the product of $P(X \leq 1)$ and $P(Y \leq 1)$.

$$P(X \leq 1 \text{ and } Y \leq 1) = 0.01 + 0.02 + 0.03 + 0.06 = 0.12$$

$$P(X \leq 1) = 0.10 + 0.20 = 0.30$$
$$P(Y \leq 1) = 0.10 + 0.30 = 0.40$$

$$P(X \leq 1)P(Y \leq 1) = 0.30 \cdot 0.40 = 0.12$$

c) Express the event that the total number of beds occupied at the two hospitals combined is at most 1 in terms of X and Y, and then calculate this probability.

$$P(X = 0 \text{ and } Y = 0 \text{ or } X = 1 \text{ and } Y = 0 \text{ or } X = 0 \text{ and } Y = 1) = 0.01 + 0.02 + 0.03 = 0.06$$

d) What is the probability that at least one of the two hospitals has no beds occupied?

$$P(X = 0 \text{ or } Y = 0) = p_X(0) + p_Y(0) - p(0,0) = 0.10 + 0.10 - 0.01 = 0.19$$

5.18

Refer to Exercise 1 and answer the following questions:

a) Given that $X = 1$, determine the conditional pmf of Y — i.e., $p_{Y|X}(0 \mid 1)$, $p_{Y|X}(1 \mid 1)$, and $p_{Y|X}(2 \mid 1)$.

y	0	1	2
$f(y \mid x = 1)$	0.2353	0.5882	0.1765

b) Given that two hoses are in use at the self-service island, what is the conditional pmf of the number of hoses in use on the full-service island?

y	0	1	2
$f(y \mid x = 2)$	0.12	0.28	0.60

c) Use the result in part (b) to calculate the conditional probability $P(Y \leq 1 \mid X = 2)$.

$$P(Y \leq 1 \mid X = 2) = 0.12 + 0.28 = 0.40$$

d) Given that two hoses are in use at the full-service island, what is the conditional pmf of the number in use at the self-service island?

x	0	1	2
$f(x \mid y = 2)$	0.0526	0.1579	0.7895

5.39

It is known that 80% of all brand A external hard drives work in a satisfactory manner throughout the warranty period (are "successes"). Suppose that $n = 15$ drives are randomly selected. Let $X =$ the number of successes in the sample. The statistic $\frac{X}{n}$ is the sample proportion (fraction) of successes. Obtain the sampling distribution of this statistic. [Hint: One possible value of $\frac{X}{n}$ is 0.2, corresponding to $X = 3$. What is the probability of this value (what kind of rv is X)?]

Since $X \sim B(15, 0.8)$, the sampling distribution for $\frac{X}{n}$ can be obtained from the table in appendix A:

x	$P(X = x)$
$\frac{0}{15}$	0.000
$\frac{1}{15}$	0.000
$\frac{2}{15}$	0.000
$\frac{3}{15}$	0.000
$\frac{4}{15}$	0.000
$\frac{5}{15}$	0.000
$\frac{6}{15}$	0.001
$\frac{7}{15}$	0.003
$\frac{8}{15}$	0.014
$\frac{9}{15}$	0.043
$\frac{10}{15}$	0.103
$\frac{11}{15}$	0.188
$\frac{12}{15}$	0.250
$\frac{13}{15}$	0.231
$\frac{14}{15}$	0.132
$\frac{15}{15}$	0.035

5.46

Young's Modulus is a quantitative measure of stiffness of an elastic material. Suppose that for aluminum alloy sheets of a particular type, its mean value and standard deviation are 70 GPa and 1.6 GPa, respectively (values given in the article "Influence of Material Properties Variability on Springback and Thinning in Sheet Stamping Process: A Stochastic Analysis" (Intl. J. of Advanced Manuf. Tech., 2010: 117–134)).

a) If $\bar{X}$ is the sample mean Young's Modulus for a random sample of $n = 16$ sheets, where is the sampling distribution of $\bar{X}$ centered, and what is the standard deviation of the $\bar{X}$ distribution?

$$\mu_{\bar{X}} = \mu = 70 \, \text{GPa}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N} = \frac{1.6^2}{16}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{1.6}{4} = 0.4 \, \text{GPa}$$

b) Answer the question posed in part (a) for a sample size of $n = 64$ sheets.

$$\mu_{\bar{X}} = \mu = 70 \text{ GPa}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N} = \frac{1.6^2}{64}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{1.6}{8} = 0.2 \text{ GPa}$$

c) For which of the two random samples, the one of part (a) or the one of part (b), is $\bar{X}$ more likely to be within 1 GPa of 70 GPa? Explain your reasoning.

For (b), $N = 64$ is sufficiently large for the CLT to apply, and thus the sampling distribution is approximately normal. Since $\sigma_{\bar{X}} = 0.2$ GPa, to be within 1 GPa of 70 GPa would mean that $\bar{X}$ is within $5\sigma_{\bar{X}}$ and thus with a probability of practically 1. So although we do not know the exact distribution for (a), the probability can be no better than (b).

5.47

Refer to Exercise 46. Suppose the distribution is normal (the cited article makes that assumption and even includes the corresponding normal density curve).

a) Calculate $P(69 \leq \bar{X} \leq 71)$ when $n = 16$.

$$P(69 \leq \bar{X} \leq 71) = P\left(\frac{69 - 70}{0.4} \leq Z \leq \frac{71 - 70}{0.4} \right)$$
$$= P(-2.50 \leq Z \leq 2.50)$$
$$= \Phi(2.50) - \Phi(-2.50)$$
$$= 0.9938 - 0.0062$$
$$= 0.9876$$

b) How likely is it that the sample mean diameter exceeds 71 when $n = 25$?

$$P(71 < \bar{X}) = 1 - P(\bar{X} \leq 71)$$
$$= 1 - P\left(Z \leq \frac{70 - 71}{\frac{1.6}{\sqrt{25}}} \right)$$
$$= 1 - P(Z \leq -3.13)$$
$$= \Phi(-3.13)$$
$$= 0.0009$$

5.49

There are 40 students in an elementary statistics class. On the basis of years of experience, the instructor knows that the time needed to grade a randomly chosen first examination paper is a random variable with an expected value of 6 min and a standard deviation of 6 min.

a) If grading times are independent and the instructor begins grading at 6:50 PM and grades continuously, what is the (approximate) probability that he is through grading before the 11:00 PM TV news begins?

Let $X = n\bar{X} =$ total time needed to grade all the tests.

$$X \approx \mathrm{N}(240, 40 \cdot 6^2)$$

The news starts in 250 minutes from the grading start time, so:

$$
\begin{aligned}
P(X \leq 250) &= P\left(Z \leq \frac{250 - 240}{6\sqrt{40}}\right) \\
&= P(Z \leq 0.26) \\
&= \Phi(0.26) \\
&= 0.6026
\end{aligned}
$$

b) If the sports report begins at 11:10, what is the probability that he misses part of the report if he waits until grading is done before turning on the TV?

$$
\begin{aligned}
P(X > 260) &= 1 - P(X \leq 260) \\
&= 1 - P\left(Z \leq \frac{260 - 240}{6\sqrt{40}}\right) \\
&= 1 - P(Z \leq 0.53) \\
&= 1 - \Phi(0.53) \\
&= 1 - 0.7019 \\
&= 0.2981
\end{aligned}
$$

5.53

Rockwell hardness of pins of a certain type is known to have a mean value of 50 and a standard deviation of 1.2.

a) If the distribution is normal, what is the probability that the sample mean hardness for a random sample of 9 pins is at least 51?

$$\bar{X} \sim \mathrm{N}(50, \frac{1.2^2}{9})$$

$$P(\bar{X} \geq 51) = 1 - P(\bar{X} \leq 51)$$
$$= 1 - P\left(Z \leq \frac{51 - 50}{\frac{1.2}{3}}\right)$$
$$= 1 - P(Z \leq 2.50)$$
$$= 1 - \Phi(2.50)$$
$$= 1 - 0.9938$$
$$= 0.0062$$

b) Without assuming population normality, what is the (approximate) probability that the sample mean hardness for a random sample for 40 pins is at least 51?

By the CLT:

$$\bar{X} \approx \mathrm{N}\left(50, \frac{1.2^2}{40}\right)$$

$$P(\bar{X} \geq 51) = 1 - P(\bar{X} \leq 51)$$
$$= 1 - P\left(Z \leq \frac{51 - 50}{\frac{1.2}{\sqrt{40}}}\right)$$
$$= 1 - P(Z \leq 5.27)$$
$$= 1 - \Phi(5.27)$$
$$= 1 - 1$$
$$= 0$$

6.7

a) A random sample of 10 houses in a particular area, each of which is heated with natural gas, is selected and the amount of gas (therms) used during the month of January is determined for each house. The resulting observations are: 103, 156, 118, 89, 125, 147, 122, 109, 138, 99. Let μ denote the average gas usage during January by all houses in this area. Compute a point estimate of μ.

$$\hat{\mu} = \bar{X} = \frac{103 + 156 + 118 + 89 + 125 + 147 + 122 + 109 + 138 + 99}{10} = 120.6 \text{ therms}$$

b) Suppose there are 10,000 houses in this area that use natural gas for heating. Let τ denote the total amount of gas used by all of these houses during January. Estimate τ using the data of part (a). What estimator did you use to compute your estimate?

$$\hat{\tau} = n\bar{X} = 10000 \cdot 120.6 = 1{,}206{,}000 \text{ therms}$$

c) Use the data in part (a) to estimate p, the proportion of all houses that used at least 100 therms.

$$\hat{p} = \frac{8}{10} = 0.80$$

d) Give a point estimate of the population median usage (the middle value in the population of all houses) based on the sample data of part (a). What estimator did you use?

First, sort the data and then use the sample median:

89,99,103,109,118,122,125,138,147,156

$$\hat{\tilde{\mu}} = \frac{118 + 122}{2} = 120.0 \text{ therms}$$

6.13

Consider a random sample $X_1, \ldots, X_n$ from the pdf:

$$f(x; \theta) = 0.5(1 + \theta x)$$

where $-1 \leq x \leq 1$ (this distribution arises in particle physics). Show that $3\hat{\bar{X}}$ is an unbiased estimator of θ. [Hint: First determine $\mu = E(X) = E(\bar{X})$.]

$$
\begin{aligned}
E(X) &= \int_{-1}^{1} x \left[\frac{1}{2}(1 + \theta x) \right] dx \\
&= \frac{1}{2} \int_{-1}^{1} (x + \theta x^2) dx \\
&= \frac{1}{2} \left[\frac{1}{2} x^2 + \frac{1}{3} \theta x^3 \right]_{-1}^{1} \\
&= \frac{1}{2} \left[\left(\frac{1}{2} + \frac{1}{3}\theta \right) - \left(\frac{1}{2} - \frac{1}{3}\theta \right) \right] \\
&= \frac{1}{2} \left(\frac{2}{3}\theta \right) \\
&= \frac{1}{3}\theta
\end{aligned}
$$

Now, we want:

$$E(X) = \hat{\mu} = \bar{X} = \frac{1}{3}\hat{\theta}$$

And therefore:

$$\hat{\theta} = 3\bar{X}$$