LAB # 5
Probability, Sampling Distribution, Central Limit Theorem

1. **Poverty in the United States**. The table below gives data on poverty in the US, from the Statistical Abstract.

	White	Black	Other	Total
Under *18* years	9752	5125	805	**15,682**
18 to *24* years	3274	1264	316	**4,854**
25 to *64* years	10261	3785	881	**14,927**
65 years and over	2939	702	114	**3,755**
Total	**26,226**	**10,876**	**2,116**	**39,218**

   a. How many people below the poverty level were there in 1993?
   39,218,000
   b. What percent of the people below the poverty level were *65* or older?
   3755 / 39218 = 0.0957 = 9.57%
   c. What percent of the people below the poverty level were black?
   10876 / 39218 = 0.2773 = 27.73%
   d. What percent of the whites below the poverty level were *65* over older?
   2939 / 26226 = 0.1121 = 11.21%
   e. Of all the children under *18* below the poverty level, what percent were black?
   5125 / 15682 = 0.3268 = 32.68%
   f. You want to know what percent of all people *65* and older were below the poverty level. Can you learn the answer from this table? Explain.
   No. We do not know the number of people who are 65 or older.
   The percentage we seek is 3755 / (# of people 65 or older)

2. At the opening of a new hit movie the age and sex of 400 moviegoers was recorded and is given in the table below.

	Age				
Gender	0 -< 13	13 -< 20	20 -< 30	Over 30	**Totals**
Female	14	72	86	22	**194**
Male	10	109	63	24	**206**
Totals	**24**	**181**	**149**	**46**	**400**

   a. Estimate the probability that a moviegoer is female.
   $$P(F)=\frac{194}{400}=0.485$$
   b. Estimate the probability that a moviegoer is younger than 30.
   $$P(\text{under } 30)=\frac{24+181+149}{400}=\frac{354}{400}=0.885$$
   c. Estimate the probability that a moviegoer is female or younger than 30.
   $$P(\text{under } 30 \text{ or } F)=\frac{354+22}{400}=0.94$$
   d. If we know a female moviegoer is selected, estimate the probability that she is a teenager.

$$P(T \mid F) = \frac{72}{194} = 0.371$$

e. If two moviegoers are selected independently of each other, estimate the probability that both are male.

$P(M) = 1 - P(F) = 0.515,$ see part a

$$P(M_1 M_2) = P(M_1)P(M_2) = 0.515(0.515) = 0.265$$

f. Is the event that a moviegoer is female independent of the event that the moviegoer is over 30? Explain.

If these two events are equal then $P(F \text{ and over } 30) = P(F)P(\text{over } 30)$**. We can check and see if this is true.**

$$P(F \text{ and over } 30) = \frac{22}{400} = 0.055$$

$$P(F) = 0.485$$

$$P(\text{over } 30) = 0.115$$

Now since 0.485 (0.115) = 0.055775 ≠ 0.055, these two events are not independent.

3. Canada has developed a new missile that hits its target 60% of the time. If two missiles are fired independently of each other at a target, what is the probability that

a. Both missiles hit the target?

$$P(H_1 \text{ and } H_2) = P(H_1)P(H_2) = 0.6 \times 0.6 = 0.36$$

b. Both missiles miss the target?

$$P(M_1 \text{ and } M_2) = P(M_1)P(M_2) = 0.4 \times 0.4 = 0.16$$

c. At least one missile hits the target?

$$P(\text{at least } 1) = 1 - P(\text{both miss}) = 1 - 0.16 = 0.84$$

d. Exactly one missile hits the target?

$$P(\text{exactly } 1) = P(\text{at least } 1) - P(\text{both hit})$$
$$= 0.84 - 0.36 = 0.48$$

4. The incubation temperature to hatch ostrich eggs for a Type SR-50 incubator is set at *99*°F and is allowed to vary with a standard deviation of *2*°F. Each hour the temperature is recorded at *50* randomly selected times. If the average of the measurements is less than *98.5*°F or greater than *99.5*°F, an alarm sounds. What is the probability that the alarm goes off?

$$P(98.5 < \bar{x} < 99.5) = P\left(\frac{98.5 - 99}{2/\sqrt{50}} < z < \frac{99.5 - 99}{2/\sqrt{50}}\right)$$
$$= P(-1.7677 < z < 1.7677) = 0.9616 - 0.0384 = 0.9232$$

The probability we seek is *1 – 0.9232 = 0.0768*

5. A college math professor has surveyed his records and found the following frequency distribution of the grades he has assigned to the 1187 students who have taken his statistics classes.

Grade	A	B	C	D	F
Frequency	125	352	461	187	62

a. Estimate the probability that a randomly selected student received an A.

$$P(A) = \frac{125}{1187} = 0.1053$$

b. Estimate the probability that a randomly selected student received an A or a B.

$$P(A \text{ or } B) = \frac{125+352}{1187} = 0.4019$$

c. What is the probability that two independently selected students both receive an A?

$$P(A_1 \text{ and } A_2) = P(A_1)P(A_2) = 0.1053(0.1053) = 0.0111$$

d. What is the probability that two independently selected students both fail?

$$P(F) = \frac{62}{1187} = 0.0522$$

$$P(F_1 \text{ and } F_2) = P(F_1)P(F_2) = 0.0522(0.0522) = 0.0027$$

6. Suppose that the data in a population follow a N(200, 5) distribution.

a. What is the probability of an individual in the population being less than 190?

$$P(x < 190) = P\left(z < \frac{190-200}{5}\right) = P(z < -2.00) = 0.0228$$

b. Suppose that the average $\bar{x}$ of a random sample of size 49 from this distribution is calculated. What is the probability that the average of a random sample of size 49 will be less than 190?

$$P(x < 190) = P\left(z < \frac{190-200}{5/\sqrt{49}}\right) = P(z < -14.00) < 0.0003$$

In other words, essentially zero

c. Find the number c so that $P(\bar{x} > c) = 0.05$

$$P(\bar{x} < c) = P\left(z < \frac{c-200}{5/\sqrt{49}}\right) = 0.95$$

$$\Rightarrow \frac{c-200}{5/\sqrt{49}} = 1.64$$

$$\Rightarrow c = 200 + 1.64 * 5/\sqrt{49} = 201.17$$

7. Assume the heights of women are Normally distributed with a mean of 64 inches (5'4") and a standard deviation of 2.5 inches. In a sample of 49 women, what is the probability that their average height would be less than 63 inches (5'3")?

$$P(\bar{x} < 63) = P\left(z < \frac{63-64}{2.5/\sqrt{49}}\right) = P(z < -2.80) = 0.0026$$

8. SAT verbal scores are normally distributed with a mean of 430 and a standard deviation of 120 (data from the College Board ATP). What score does a class of 15 students need to average in order to be in the top 10%?

$$P(\bar{x} < c) = P\left(z < \frac{c-430}{120/\sqrt{15}}\right) = 0.90$$

$$\Rightarrow \frac{c-430}{120/\sqrt{15}} = 1.28$$

$$\Rightarrow c = 430 + 1.28 * 120/\sqrt{15} = 469.66$$

9. The diameters of bolts produced by a certain machine are normally distributed with a mean of 0.30 inches and a standard deviation of 0.01 inches. What percentage of bolts will have a diameter greater than 0.32 inches?

$$P(x > 0.32) = P\left(Z > \frac{0.32-0.30}{0.015}\right) = P(Z > 1.33) = 1 - 0.9082 = 0.0918$$

10. Let x_1, x_2, …, x_{100} denote the actual net weights (in pounds) of *100* selected bags of fertilizer. Suppose that the weight of a randomly selected bag is a random variable that has a mean of $\mu = 50$ pounds and a standard deviation of $\sigma = 2$ pounds. Let $\bar{x}$ be the sample mean weight.

   a. Determine the probability that the sample mean is between *49.50* and *50.50* pounds?

$$P(49.5 < \bar{x} < 50.5) = P\left(\frac{49.5-50}{0.2} < z < \frac{50.50-50}{0.2}\right)$$

$$= P(-2.5 < z < 2.5) = 0.9938 - 0.0062 = 0.9876$$

   b. Determine the probability that a single bag of fertilizer will be between *49.50* and *50.50*?

$$P(49.5 < x < 50.5) = P\left(\frac{49.5-50}{2.0} < z < \frac{50.50-50}{2.0}\right)$$

$$= P(-0.25 < z < 0.25) = 0.5987 - 0.4013 = 0.1974$$

11. Water permeability of concrete is an important characteristic in assessing suitability for various applications. Permeability can be measured by letting water flow across the surface and determining the amount lost (inches/hour). Suppose that the permeability index *X* for a randomly selected concrete specimen of a particular type has a normal distribution with $\mu = 1000$ and $\sigma = 150$.

   a. How likely is it that a single specimen will have a permeability index between *850* and *1300*?

$$P(850 < x < 1300) = P\left(\frac{850-1000}{150} < z < \frac{1300-1000}{150}\right)$$

$$= P(-1.00 < z < 2.00) = 0.9772 - 0.1587 = 0.8185$$

   b. How likely is it that the average of four specimens will have a permeability index between *850* and *1300*?

$$P(850 < x < 1300) = P\left(\frac{850-1000}{150/\sqrt{4}} < z < \frac{1300-1000}{150/\sqrt{4}}\right)$$

$$= P(-2.00 < z < 4.00) = 1.0000 - 0.0228 = 0.9772$$

c. How likely is it that the average of forty specimens will have a permeability index between *850* and *1300*?

$$P(850 < x < 1300) = P\left(\frac{850-1000}{150/\sqrt{40}} < z < \frac{1300-1000}{150/\sqrt{40}}\right)$$

$$= P(-6.32 < z < 12.65) = 1 - 0 = 1$$