UNIVERSITY OF JORDAN
CHEMICAL ENGINEERING DEPARTMENT
0905331 –PROCESS MODELING BY STATISTICAL METHODS

1. **(30 points)** A batch contains 36 bacteria cells and 12 of the cells are not capable of cellular replication. Suppose you examine three bacteria cells selected at random, without replacement.

   a. $f(x) = \dfrac{\binom{24}{x}\binom{12}{3-x}}{\binom{36}{3}}$

   b. $\mu = E(X) = np = (3)(24)/36 = 2$

   $V(X) = np(1-p)\dfrac{N-n}{N-1} = 0.629.$

   c. $P(X \le 2) = 1 - P(X = 3) = 0.717.$

2. **(30 points)** Natural red hair consists of two genes. People with red hair have two dominant genes, two regressive genes, r one dominant and one regressive gene. A group of 1000 people was categorized as follows:

	Gene 2		
Gene 1	**Dominant**	**Regressive**	**Other**
Dominant	5	25	30
Regressive	7	63	35
Other	20	15	800

Let A denote the event that a person has a dominant red hair gene and let B denote the event that a person has a regressive red hair gene. If a person is selected at random from this group, compute the following.

(a) $P(A) = \dfrac{5+25+30+7+20}{1000} = 0.087$

(b) $P(A \cap B) = \dfrac{25+7}{1000} = 0.032$

(c) $P(A \cup B) = 1 - \dfrac{800}{1000} = 0.20$

(d) $P(A' \cap B) = \dfrac{63+35+15}{1000} = 0.113$

(e) $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.032}{(25+63+15+7+35)/1000} = 0.2207$

3. **(30 points)** The time it takes a cell to divide (called mitosis) is normally distributed with an average time of one hour and a standard deviation of 5 minutes.

   a. $P(X < 45) = P(Z < \dfrac{45-60}{5}) = P(Z < -3) = 0.00135.$

   b. $P(X > 65) = P(Z > \dfrac{65-60}{5}) = P(Z > 1) = 1 - P(Z < 1)$

   $= 1 - 0.841345 = 0.158655.$

   c. $P(X < x) = P(Z < \dfrac{x-60}{5}) = 0.99$.

   $Z = 2.33 \Rightarrow X = 72$ minutes.

4. **(30 points)** In a healthy individual age 20 to 29 years, the calcium level in the blood, X, is usually between 8.5 and 10.5 mg/dL and the cholesterol level, Y, and is usually between 120 and 240 mg/dL. Assume that for a healthy individual in this age group the random variable (X, Y) is uniformly distributed such that

$$f_{XY} = c \quad 8.5 \le x \le 10.5$$
$$120 \le y \le 240$$

   a. What is the value of the constant c that will make the function a mass density function?

$$\int_{8.5}^{10.5}\int_{120}^{240} c\,dy\,dx = 1$$

$$c\int_{8.5}^{10.5}(240-120)\,dx = 1$$

$$120c(10.5-8.5) = 240c = 1 \Rightarrow c = 1/240.$$

   b. Determine the probability that an individual's calcium level will lie between 9 and 10 mg/dL, whereas the cholesterol level is between 125 and 140 mg/dL.

$$P(9 \le x \le 10; 125 \le y \le 140) = \int_{9}^{10}\int_{125}^{140}(1/240)\,dy\,dx$$
$$= (1/240)\int_{9}^{10}(140-125)\,dx$$
$$= 15/240.$$

   c. Determine Cov(X, Y).

$$E(X) = \int_{8.5}^{10.5}\int_{120}^{240} x(1/240)\,dy\,dx = \int_{8.5}^{10.5} x(1/2)\,dx = x^2/4\Big|_{8.5}^{10.5} = 9.5 \text{ mg/dL}.$$

$$E(Y) = \int_{8.5}^{10.5}\int_{120}^{240} y(1/240)\,dy\,dx = (1/240)\int_{8.5}^{10.5} y^2/2\Big|_{120}^{240}\,dx$$
$$= (1/240)\int_{8.5}^{10.5} 21{,}600\,dx = 180 \text{ mg/dL}.$$

$$E(XY) = \int_{8.5}^{10.5}\int_{120}^{240} xy(1/240)\,dy\,dx = (1/240)\int_{8.5}^{10.5} xy^2/2\Big|_{120}^{240}\,dx$$
$$= (1/240)\int_{8.5}^{10.5} 21{,}600x\,dx = (21{,}600/240)x^2/2\Big|_{8.5}^{10.5} = 1710 \text{ (mg/dL)}^2.$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 1710 - (9.5)(180) = 0!$$

   The covariance is zero which implies that knowledge that X assumes a value above its mean gives us no indications as the value of Y to its mean.

   d. Determine if X and Y are independent. The two variables are independent because E(XY) = E(X)E(Y). Don't jump to conclusions since Cov(X ,Y) = 0. Covariance can be zero for many instances other than the equality constraint!

   e. Determine ρ_{XY}. The correlation coefficient is defined as: $\rho_{XY} = \dfrac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$.

   Again this is a fool's trap if you calculate the variances of X and Y. Since the Cov(X, Y) in the numerator is zero you don't need to calculate any thing. All you need to do is to state that the correlation coefficient is zero!