

Homework #2

Instructor: Dr. Zafeirakis Zafeirakopoulos Name: Ahmed Semih Özmekik Student Id: 171044039
 Assistant: Gizem Süngü

Course Policy: Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- It is not a group homework. Do not share your answers to anyone in any circumstance. Any cheating means at least -100 for both sides.
- Do not take any information from Internet.
- No late homework will be accepted.
- For any questions about the homework, send an email to gizemsungu@gtu.edu.tr.
- Submit your homework (both your latex and pdf files in a zip file) into the course page of Moodle.
- Save your latex, pdf and zip files as "Name_Surname_StudentId".{tex, pdf, zip}.
- The deadline of the homework is 26/04/20 23:55.

Problem 1:

(10 points)

Suppose that you are inspecting a lot of 1000 surgical masks, among which 20 are used. You choose two masks randomly from the lot without replacement. Let

$$X_1 = \begin{cases} 1, & \text{if the 1st surgical mask is used,} \\ 0, & \text{otherwise} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{if the 2nd surgical mask is used,} \\ 0, & \text{otherwise} \end{cases}$$

Find the probability that at least one surgical mask chosen is used.

(Solution)

The probability we concern is an example of discrete probability distribution that describes the probability of k successes (random draws for which the object drawn has a specified feature) in n draws; which is hypergeometric distribution.

A random variable X follows the hypergeometric distribution:

$$p_X(k) = (X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

where

- N is the population size,
- K is the number of success states in the population,
- n is the number of draws (i.e. quantity drawn in each trial),
- k is the number of observed successes,
- $\binom{a}{b}$ is a binomial coefficient.

After defining our problem, since X_1 and X_2 are disjoint, we need to find the probability of each, and then add them up. Here is our substitution:

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

$$p_X(k) = \frac{\binom{20}{x} \binom{980}{2-x}}{\binom{1000}{2}}$$

Let's find $p_X(X_1)$ and $p_X(X_2)$:

$$p_X(1) = \frac{\binom{20}{1} \binom{980}{1}}{\binom{1000}{2}}$$

$$p_X(1) = 0.03923923923923924$$

$$p_X(2) = \frac{\binom{20}{2}}{\binom{1000}{2}}$$

$$p_X(2) = 0.0003803803803803804$$

$$p_X(1) + p_X(2) = 0.03961961961961962$$

Problem 2:

(8+8=16 points)

Suppose X and Y are random variables with $P(X = 1) = P(X = -1) = \frac{1}{2}$; $P(Y = 1) = P(Y = -1) = \frac{1}{2}$. Let $c = P(X = 1 \text{ and } Y = 1)$.

(a) Determine the joint distribution of X and Y , $\text{Cov}(X, Y)$, and $\text{Cor}(X, Y)$.

(Solution)

We need to find the partial function $F(x, y)$. Regarding to definition given above, we know:

$$P(X = 1) = P(X = -1) = \frac{1}{2}$$

$$P(X = 1) + P(X = -1) = 1$$

$$P(X = x_1) = 0$$

where $x_1 = 1$, or $x_1 = -1$.

Hence,

$$f(x, -1) = F(1, -1) + F(-1, -1) = \frac{1}{2}$$

$$F(x, 1) = F(1, 1) + F(-1, 1) = \frac{1}{2}$$

$$F(-1, y) = F(-1, -1) + F(-1, 1) = \frac{1}{2}$$

$$F(1, y) = F(1, -1) + F(1, 1) = \frac{1}{2}$$

$$F(-1, -1) = c$$

$$F(-1, 1) = \frac{1}{2} - c$$

$$F(1, 1) = \frac{1}{2} - c$$

$$F(1, -1) = c.$$

$$F(x, y) = \begin{cases} c, & x = -1, y = -1, \\ \frac{1}{2} - c, & x = -1, y = 1 \\ \frac{1}{2} - c, & x = 1, y = -1, \\ c, & x = 1, y = 1, \end{cases}$$

Now let's find value of $\text{Cov}(X, Y)$. The sample covariance matrix is a K-by-K matrix. $Q = [q_{jk}]$ And it can be shown as, $q_{jk} = \frac{1}{N-1} \sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$.

$$\text{Or, } \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_x \mu_y.$$

$$\begin{aligned} E(XY) &= \sum_i \sum_j x_i y_i F_{xy}(x_i, y_i) \\ &= \sum_i \sum_j x_i y_i F_x(x_i) F_y(y_j) \\ &= \left(\sum_i x_i F_x(x_i) \right) \left(\sum_j y_j F_y(y_j) \right) \\ &= E(X)E(Y) \\ &= \sum_x x \left(\sum_y F(x, y) y \right) \\ &= \sum_x (-f(x, -1) + f(x, 1)) \\ &= \sum_x (-f(x, -1) + f(x, 1)) \\ &= f(-1, -1) - f(-1, 1) - f(1, -1) + f(1, 1) \end{aligned}$$

$$\begin{aligned} \mu_x &= \sum_x g(x) x \\ &= -1 \frac{1}{2} + 1 \frac{1}{2} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \mu_y &= \sum_y h(y) y \\ &= -1 \frac{1}{2} + 1 \frac{1}{2} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY) - \mu_x \mu_y \\ &= c - \frac{1}{2} + c - \frac{1}{2} + c + c \\ &= 4c - 1 \end{aligned}$$

$$\begin{aligned}
 \text{Cor}(X, Y) &= \frac{\sigma_{xy}}{\sigma_x \sigma_y} \\
 \sigma_x^2 &= E(X^2) - \mu_x^2 = E(X^2) \\
 E(X^2) &= \sum_x x^2 f_1(x) = f_1(1) + f_1(-1) \\
 &= \frac{1}{2} + \frac{1}{2} = 1 \\
 \sigma_y^2 &= E(Y^2) - \mu_y^2 = E(Y^2) \\
 E(Y^2) &= \sum_y y^2 f_2(y) = f_2(1) + f_2(-1) \\
 &= \frac{1}{2} + \frac{1}{2} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Cor}(X, Y) &= \frac{\sigma_{xy}}{\sigma_x \sigma_y} \\
 &= 4c - 1
 \end{aligned}$$

(b) For what value(s) of c are X and Y independent? For what value(s) of c are X and Y 100% correlated?

(Solution)

Independence: $F(x, y) = f_1(x)f_2(y)$. Remember our equation from above:

$$\begin{aligned}
 E(X, Y) &= \sum_x \left(\sum_y F(x, y)xy \right) \\
 &= \sum_x x f_1(x) \sum_y y f_2(y) \\
 &= (f_1(-1) + f_1(1)) \times (f_2(-1) + f_2(1)) \\
 &= \left(\frac{1}{2} - \frac{1}{2}\right) \times \left(\frac{1}{2} - \frac{1}{2}\right) \\
 &= 0 \\
 E(x, y) &= 0 \\
 4c - 1 &= 0 \\
 c &= \frac{1}{4}
 \end{aligned}$$

To satisfy %100 correlation:

$$\begin{aligned}
 |\text{Cor}(X, Y)| &= 1 \\
 |4c - 1| &= 0 \\
 c_1 &= \frac{1}{2} \\
 c_2 &= 0
 \end{aligned}$$

Problem 3:

(4+4+4+4+4+4=24 points)

In the information security department in a software company, a single crucial program works only 85% of the time. In order to enhance the reliability of the system, it is decided that 3 programs will be installed in parallel such that the system fails only if they all fail. Assume the programs act independently and that they are equivalent in the sense that all 3 of them have an 85% success rate. Consider the random variable X as the number of components out of 3 that fail.

(a) Write out a probability function for the random variable X .

(Solution)

In the theory of probability and statistics, a Bernoulli trial (or binomial trial) is a random experiment with exactly two possible outcomes, "success" and "failure", in which the probability of success is the same every time the experiment is conducted.

The probability of exactly k successes in the experiment $B(n, p)$ is given by:

$$P(k) = \binom{n}{k} p^k q^{n-k}$$

Let's substitute.

$$\begin{aligned} P(x) &= B(x; n, p) = b(x, 3, 15/100) \\ &= \binom{3}{x} \left(\frac{15}{100}\right)^x \left(\frac{85}{100}\right)^{3-x} \end{aligned}$$

(b) What is $E(X)$ (i.e., the mean number of programs out of 3 that fail)?

(Solution)

$$\begin{aligned} \mu &= np \\ \mu &= 3\left(\frac{15}{100}\right) \\ &= \frac{45}{100} \end{aligned}$$

(c) What is $\text{Var}(X)$?

(Solution)

$$\begin{aligned} \text{Var}(x) &= np(1-p) \\ &= 3 \frac{15}{100} \frac{85}{100} \\ &= 0.3825 \end{aligned}$$

(d) What is the probability that the entire system is successful?

(Solution)

For success, in 3 attempts, there can be fails as the following 0, 1, 2. Meaning that:

$$\begin{aligned} f(0) + f(1) + f(2) &= 1 - f(3) \\ &= 1 - \binom{3}{3} \left(\frac{15}{100}\right)^3 \\ &= 0.996625 \end{aligned}$$

(e) What is the probability that the system fails?

(Solution)

For fail, in 3 attempts, 3 tests should fail. Meaning that:

$$\begin{aligned} f(3) &= \binom{3}{3} \frac{15}{100} \\ &= 0.003375 \end{aligned}$$

(f) If the desire is to have the system be successful with probability 0.99, are three programs sufficient? If not, how many are required?

(Solution)

0.996625 is greater than 0.99. Hence, it is sufficient.

Problem 4:

(10+10=20 points)

According to World Health Organization (WHO), approximately 30% of all treatment failures in Covid-19 are caused by lack of available respirators.

(a) What is the probability that out of the next 20 treatment failures at least 10 are due to lack of available respirators?

(Solution)

Again, the probability of exactly k successes in the experiment $B(n, p)$ is given by:

$$\begin{aligned} P(k) &= b(x, n, p) = \binom{n}{k} p^k q^{n-k} \\ &= \binom{20}{x} \left(\frac{3}{10}\right)^x \left(\frac{7}{10}\right)^{20-x} \end{aligned}$$

We will sum up the probabilities from $x = 10$ to $x = 20$ as stated in the question.

$$\begin{aligned} P &= \sum_{x=10}^{20} f(x) \\ &= \binom{20}{10} \left(\frac{3}{10}\right)^{10} \left(\frac{7}{10}\right)^{20-10} + \binom{20}{11} \left(\frac{3}{10}\right)^{11} \left(\frac{7}{10}\right)^{20-11} + \dots \\ &= 0.04796189733 \end{aligned}$$

(b) What is the probability that no more than 4 out of 20 such failures are due to lack of available respirators?

(Solution)

Again, for this question, only our boundaries will change. Hence

$$\begin{aligned} P &= \sum_{x=0}^4 f(x) \\ &= \binom{20}{0} \left(\frac{3}{10}\right)^0 \left(\frac{7}{10}\right)^{20-0} + \dots \\ &= 0.23750777887760 \end{aligned}$$

Problem 5:

(7+7=14 points)

A manufacturing company uses an acceptance scheme on items from a production line before they are shipped. The plan is a two-stage one. Boxes of 25 items are ready for shipment, and a sample of 3 items is tested for defectives. If any defectives are found, the entire box is sent back for 100% screening. If no defectives are found, the box is shipped.

(a) What is the probability that a box containing 3 defectives will be shipped?

(Solution)

We can determine the probability with a simple count. Once

$$\begin{aligned} P &= \frac{\binom{22}{3}}{\binom{25}{3}} \\ &= \frac{77}{115} \end{aligned}$$

(b) What is the probability that a box containing only 1 defective will be sent back for screening?

(Solution)

First of all, as in any case, we will pick 3 out of 25, $\binom{25}{3}$. Then, from this space we know that there is only 1 defective and we are looking for this defective item to appear. Of course, now we know that there are 24 items left which are non-defective. So we have 3 test. For 2 of them we pick $\binom{24}{2}$. Meaning that, we make sure that for 3 tests we picked 2 as non-defective. And we have only 1 defective left, if we are to pick this, we could've picked it as $\binom{1}{1}$, but since there is only left, we have no choice but to pick the only one. So this makes the probability:

$$\begin{aligned} P &= \frac{\binom{24}{2}}{\binom{25}{3}} \\ &= \frac{3}{25} \\ &= 0.12 \end{aligned}$$

Problem 6: Probability Distributions of Random Variables

(8+8=16 points)

Suppose the probability that any given person will believe a tale about the transgressions of a famous actress is 0.8. What is the probability that

(a) the sixth person to hear this tale is the fourth one to believe it?

(Solution)

We want the sixth person to hear it to be the fourth person to believe it. So 3 person will believe this tale. But 6 person will hear it. Let's go by one up to to 5th person, but of course we need to handle the orders. Because of that, we pick 3 persons out of 5, as a believer with $\binom{5}{3}$. And multiply this by with the believing probability (for each), which is $\left(\frac{3}{5}\right)\left(\frac{8}{10}\right)^3$. And now we skip two people by picking them as an non-believer. To do this, let's multiply it with $\left(\frac{2}{10}\right)^2$:

$$= \binom{3}{5} \left(\frac{8}{10}\right)^3 \left(\frac{2}{10}\right)^2$$

This is just before the last believer, which is 4th believer, and the 6th person:

$$\begin{aligned}
 &= \binom{3}{5} \left(\frac{8}{10}\right)^3 \left(\frac{2}{10}\right)^2 \frac{8}{10} \\
 &= \binom{3}{5} \left(\frac{8}{10}\right)^4 \left(\frac{2}{10}\right)^2 \\
 &= 0.16384
 \end{aligned}$$

(b) the third person to hear this tale is the first one to believe it?

(Solution)

First two person will not believe it. Again probability of not believing is 0.2, Probability of the the two person will not believe is, $(\frac{2}{10})^2$. And then we multiply this by 0.8, since the third person will believe it:

$$\begin{aligned}
 P &= \left(\frac{2}{10}\right)^2 \frac{8}{10} \\
 &= \frac{4}{125} \\
 &= 0.032
 \end{aligned}$$