Unit 4: Probability Distribution

Lesson 4.3: Hypergeometric distribution

<u>Minds on</u>: On a team of 15 astronauts, 6 are women and 9 are men. If four astronauts are selected at random for a flight simulation, what is the probability that two men and two women are selected?

A **<u>hypergeometric distribution</u>** has a specific number of **dependent** trials with two possible outcomes, **success** or **failure**.

The probability of success changes as each trial is made since each selection reduces the number of items that could be selected in the next trial.

The random variable is the number of successful outcomes in a specified number of trials.

Probability in a Hypergeometric Distribution:
$P(x) = \frac{{}_aC_x \times {}_{n-a}C_{r-x}}{{}_nC_r}$, where ***a*** is the number of available outcomes among a total of ***n*** possible outcomes in ***r*** trials.

Expectation for a Hypergeometric Distribution:
The expectation for a hypergeometric distribution is, $E(x) = \frac{ra}{n}$

<u>Example 1</u>: Back to Minds-on question, define *n, a, r*, and *x*, and what is the expected number of women to be selected for the simulation mission?

Example 2: Determine the probability distribution for the number of women on a civil-court jury (6 jurors) selected from a pool of 8 men and 10 women?

The selection process is __________________ since each person who is chosen for the jury cannot be selected again.

# women, x	Probability, P(x)

What is the expected number of women on the jury?

Example 3: In the spring, the Ministry of the Environment caught and tagged 500 raccoons in a wilderness area. The raccoons were released after being vaccinated against rabies. To estimate the raccoon population in the area, the ministry caught 40 raccoons during the summer. Of these, 15 had tags.

a) Determine whether this situation can be modeled with a hypergeometric distribution.

Since the 40 racoons are all different (no repetition) the trails are __________________.

Racoons are either tagged (______________) or not (_______________).

Therefore, ____.

b) Assuming that the number of tagged raccoons caught during the summer is equal to the expectation for the hypergeometric distribution, estimate the raccoon population in the wilderness area.

Homework: A box contains seven yellow, three green, five purple, and six red Smarties jumbled together.

a) What is the expected number of **red** Smarties among **five** Smarties poured from the box?

b) Verify that the expectation formula for a hypergeometric distribution gives the same result as the general equation for the expectation of any probability distribution.

c) What is the probability of **at least** two purple Smarties in the box?

ANSWER

Date:

A box contains seven yellow, three green, five purple, and six red Smarties™ jumbled together.

a) What is the expected number of **red** Smarties among **five** Smarties poured from the box?

$n = 7+3+5+6 = 21$

$r = 5$ chosen

$a = 6$ red

$E(x) = \frac{ra}{n} = \frac{5(6)}{21}$

$= \frac{30}{21} = 1.43$

∴ You can expect on average to get 1.5 reds in a group of 5

b) Verify that the expectation formula for a hypergeometric distribution gives the same result as the general equation for the expectation of any probability distribution

$E(x) = \Sigma x P(x)$

$= 0\left(\frac{{}_6C_0 \cdot {}_{15}C_5}{{}_{21}C_5}\right) + 1\left(\frac{{}_6C_1 \cdot {}_{15}C_4}{{}_{21}C_5}\right) + 2\left(\frac{{}_6C_2 \cdot {}_{15}C_3}{{}_{21}C_5}\right) + 3\left(\frac{{}_6C_3 \cdot {}_{15}C_2}{{}_{21}C_5}\right) +$

$4\left(\frac{{}_6C_4 \cdot {}_{15}C_1}{{}_{21}C_5}\right) + 5\left(\frac{{}_6C_5 \cdot {}_{15}C_0}{{}_{21}C_5}\right)$

$= \frac{1}{{}_{21}C_5}(8190 + 13650 + 6300 + 900 + 30)$

$= \frac{29070}{20349} = 1.43$

we see both forms of the expected value are the same.

c) What is the probability of ***at least*** two purple Smarties in the box?

$P(x \geq 2) = 1 - P(0) - P(1)$ and $P(x) = \frac{{}_aC_x \cdot {}_{n-a}C_{r-x}}{{}_nC_r}$

$n = 21$

$r = 5$

$a = 5$

$x = 0, 1$

$= 1 - \frac{{}_5C_0 \cdot {}_{16}C_5}{{}_{21}C_5} - \frac{{}_5C_1 \cdot {}_{16}C_4}{{}_{21}C_5}$

$= 1 - 0.2147 - 0.4472$

$= 0.3381$

$\approx 33.8\%$

∴ there is a 33.8% chance of getting 2 purple Smarties in a sample of 5 smarties.