NAME: solutions

TEST 2

Introduction to Statistical Methods
(Analytical Chemistry)
201-922-DW
Instructor: Emilie Richer
November 7, 2016

Instructions:

- You have 75 minutes to complete the test
- No books, cell phones or other communication devices are permitted
- You must show all of your work in order to be credited with full marks
- Anyone suspected of cheating will be asked to leave
- This test is marked out of **35 marks**

[QUESTION 1] (5 MARKS)

A randomly selected sample of 40 Dawson students are asked to estimate the number of hours they spend doing homework in a week. The answers are given as follows:

1	0	5	10	20	11	7	8	2	3
0	1	1	5	5	10	11	12	13	7
7	7	8	10	12	15	16	14	22	15
11	13	15	17	9	8	7	5	6	2

(a) Compute the sample mean $\bar{x}$ for the above data

$\bar{x} = 8.775$ hrs

(b) Compute the sample standard deviation S for the above data

$s = 5.51$ hrs

(c) If we selected one of the 40 surveyed Dawson students at random, what is the probability that he/she estimated doing less homework than the sample mean $\bar{x}$?

$P(x < 8.775) = \frac{21}{40}$

(count the # of students who estimated 8 hrs or less)

[QUESTION 2] (5 MARKS)

A small pond contains a total of 12 goldfish. Of these 12 goldfish, 5 have been tagged by a team of marine biologists in order to track their movement. The team randomly catches 3 goldfish (at once) from the pond for a study.

We define the discrete random variable X to be the # of tagged fish among the 3 fish caught by the team.

(a) Give the probability distribution for the discrete random variable X

$X = x$	$P(X = x)$
0	$[{}_5C_0 \cdot {}_7C_3]/{}_{12}C_3 = 35/220$
1	$[{}_5C_1 \cdot {}_7C_2]/{}_{12}C_3 = 105/220$
2	$[{}_5C_2 \cdot {}_7C_1]/{}_{12}C_3 = 70/220$
3	$[{}_5C_3 \cdot {}_7C_0]/{}_{12}C_3 = 10/220$

(b) Which two conditions do your distribution in (a) satisfy in order to be a **probability** distribution?

- all values of $P(x)$ are ≥ 0
- $\sum P(x) = 1 \quad \left(\frac{35}{220} + \frac{105}{220} + \frac{70}{220} + \frac{10}{220} = \frac{220}{220} = 1\right)$

(c) Compute the expected value μ for the probability distribution in (a).

$$\mu = \sum_{x=0}^{3} x\,P(x) = 0 \cdot \frac{35}{220} + 1 \cdot \frac{105}{220} + 2 \cdot \frac{70}{220} + 3 \cdot \frac{10}{220}$$

$$= \frac{275}{220} = 1.25$$

(d) Compute the probability that the research team catches more tagged fish than expected?

$$P(x > 1.25) = P(2) + P(3) = 70/220 + 10/220 = 80/220 = 4/11 \quad (0.364)$$

[QUESTION 3] (5 MARKS)

A coach claims that his team has a 65% chance of winning each of its upcoming games. Given that the coach is correct,

(a) What is the probability that his team will win at least 2 of its next 7 games?

BINOMIAL WITH $n = 7$ $p = 0.65$ $q = 0.35$

x = # of WINS AMONG NEXT 7 GAMES

WE WANT: $P(X \geq 2)$

$$P(X = 1) = {}_7C_1 (0.65)^1 \cdot (0.35)^6$$

$$= 0.0466$$

$$P(X = 0) = {}_7C_0 (0.65)^0 (0.35)^7 = 0.000643$$

$$P(X \geq 2) = 1 - [P(0) + P(1)] = 1 - [0.009008]$$

$$= 0.99099$$

THE PROBABILITY IS 0.991 THAT THE TEAM WINS AT LEAST 2 of the NEXT 7 GAMES

(b) What is the probability that the team must play 10 games before getting its next 3 wins and that the 3rd win occurs on the 10th game?

WE NEED TWO EVENTS TO HAPPEN

A: WINNING 2 GAMES IN 9 GAMES

B: WINNING THE ~~3rd GAME ON THE~~ 10th GAME

A is a BINOMIAL WITH $n = 9$ $p = 0.65$

WE HAVE $P(A) = P(X = 2) = {}_9C_2 (0.65)^2 (0.35)^7$

$$= 0.009786$$

$$P(\text{3rd WIN ON 10th GAME}) = P(A \cap B) = P(A) \cdot P(B)$$

$$= (0.009786) \cdot (0.65)$$

$$= 0.00636$$

THE PROB. OF WINNING 3rd WIN ON 10th GAME IS 0.00636.

[QUESTION 4] (5 MARKS)
Consider the following set of sample data given as a frequency table:

Value	Frequency
4	1
2	7
0	5
-1	12

For the data set, using your calculator, compute:

a) $\sum(x_i)$ = 6

b) $\sum(x_i)^2$ = 56

c) the sample mean $\overline{X}$ = 0.24

d) the sample standard deviation S_x = 1.5078

e) the total variation SS(X) = ~~88.66~~ 54.56

[QUESTION 5] (5 MARKS)
Six cards are selected at random from a standard 52-card deck of playing cards. Compute the probability of having selected the following type of 6-card hands:

(a) A hand with only 2 red cards

$$\frac{({}_{26}C_2)\cdot({}_{26}C_4)}{{}_{52}C_6} \quad (= 0.2387)$$

(b) A hand with at most 1 spade

$$\cancel{1-P} \quad P(0\text{ spades}) + P(1\text{ spade})$$
$$= \frac{{}_{39}C_{6}}{{}_{52}C_6} + \frac{({}_{13}C_1)\cdot({}_{39}C_5)}{{}_{52}C_6} \quad (= 0.\cancel{3677}\ 5279)$$

(c) A hand with all diamonds

$$\frac{{}_{13}C_6}{{}_{52}C_6} \quad (= 0.0000843)$$

(d) A hand with no even faces (WHAT I MEANT HERE IS NO 2, 4, 6, 8, 10)

$$\frac{{}_{32}C_6}{{}_{52}C_6} = (= 0.0445)$$

[QUESTION 6] (5 MARKS)
A special coin is created in such a way that when the coin is flipped it lands on "tails" 75% of the time. A random experiment consists of flipping this special coin 4 times. We define the random variable X to be the number of tails observed in the random experiment.

(a) Give the probability distribution for X

$X = x$	$P(X = x)$
0	${}_4C_0 \cdot (0.75)^0 (0.25)^4 = \frac{1}{256}$
1	${}_4C_1 \cdot (0.75)^1 (0.25)^3 = \frac{12}{256}$
2	${}_4C_2 \cdot (0.75)^2 (0.25)^2 = \frac{54}{256}$
3	${}_4C_3 \cdot (0.75)^3 (0.25)^1 = \frac{108}{256}$
4	${}_4C_4 \ (0.75)^4 (0.25)^0 = \frac{81}{256}$

Binomial $n=4$
$p = 0.75$ $q = 0.25$

(b) What type of probability distribution is illustrated in this problem? Explain.

This is a Binomial experiment:

- → n repeated, identical, independent trials
- → Two possible outcomes: success (tails) & failure (heads)
- → Probability of success $p = 1 - q$ where q = probability of failure

(c) Compute μ and σ for the probability distribution in (a)

For a Binomial distribution

$\mu = np = 4 \cdot 0.75 = 3$

$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{4 \cdot (0.75) \cdot (0.25)} = 0.866$

[QUESTION 7] (5 MARKS)
Explain the difference between $\overline{X}$ and μ and the difference between S and σ.

$\bar{x}$ & μ represent means of a sample & population respectively, they are computed in the same way: the sum of data points divided by the number of data points

s & σ represent the standard deviation of a sample & of a population, these are computed differently:

$$s = \sqrt{\frac{\text{total variation of sample}}{\text{sample size} - 1}}$$

$$\sigma = \sqrt{\frac{\text{total variation of pop.}}{\text{population size}}}$$