

STAT 260 Spring 2023: Assignment 2

Due: Friday January 27th BEFORE 11:59pm PT to Crowdmark

Please read the instructions below and in the Written Assignment 2 assignment on Crowdmark.

Solutions are to be uploaded to Crowdmark. Here you will be asked to upload your solutions to each question separately. You may hand write your solution on a piece of paper, or electronically using a tablet, or you may type your solutions. If you wish to use this question sheet and write your solutions on the page, space has been provided below. One of the quickest ways to upload handwritten work is by accessing Crowdmark from within a web browser on a smartphone. In the area where you upload work, press the "+" button. This will give you the option of using a file already on your phone, or you can use the phone camera to photograph your work. Using the phone's camera allows you to bypass the step where you need to scan your work!

For full marks, your work must be neatly written, and contain enough detail that it is clear how you arrived at your solutions. Messy, poorly formatted work will receive deductions, or may not be graded at all.

Talking to your classmates about assigned work is a healthy practice that is encouraged. However, in the end, each person is expected to write their own solutions, in their own words, and in a way that reflects their own understanding. Assignment questions are not to be posted to homework "help" websites.

Late policy: Late assignments will be accepted until the final cutoff of 11:59pm on Sunday January 29th. Solutions submitted within 1 hour of the Friday deadline will have a 5% late penalty automatically applied within Crowdmark. Solutions submitted after 1 hour of the Friday deadline but before the final Sunday cutoff will have a 20% late penalty applied. Solutions submitted after the final Sunday cutoff will be graded for feedback, but marks will not be awarded.

1. [2 marks] A factory wants to determine the relationship between the daily midday temperature (measured in degrees Celsius) and the number of defective parts produced during that day. Here x_i represents the temperature in degrees Celsius and y_i represents the number of defective parts for day i .

Day	x	y
1	30.5	36
2	26.6	24
3	25.3	13
4	28.1	35
5	25.7	19
6	24.0	22

Calculate the correlation coefficient, r , for x and y .

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$
30.5	36	$30.5 - 26.7 = 3.8$	$36 - 149/6 = 67/6 = 11.167$
26.6	24	$26.6 - 26.7 = -0.1$	$24 - 149/6 = -5/6 = -0.833$
25.3	13	$25.3 - 26.7 = -1.4$	$13 - 149/6 = -71/6 = -11.833$
28.1	35	$28.1 - 26.7 = 1.4$	$35 - 149/6 = 61/6 = 10.167$
25.7	19	$25.7 - 26.7 = -1$	$19 - 149/6 = -35/6 = -5.833$
24.0	22	$24.0 - 26.7 = -2.7$	$22 - 149/6 = -17/6 = -2.833$

$$\begin{aligned} \bar{x} &= 26.7 & \bar{y} &= \frac{149}{6} = 24.833 & s_x &= 2.3091 & s_y &= 9.0646 \\ s_{xy} &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{(3.8)(67/6) + (-0.1)(-5/6) + (-1.4)(-71/6) + (1.4)(61/6) + (-1)(-35/6) + (-2.7)(-17/6)}{6-1} \\ &= \frac{86.8}{5} = 17.36 \\ r &= \frac{s_{xy}}{s_x s_y} = \frac{17.36}{(2.3091)(9.0646)} = \boxed{0.8294} \end{aligned}$$

2. [2 marks] Students registered in the faculties of Science, Business, and Humanities were surveyed and were asked if they regularly drink coffee or tea (or both or neither). The relative frequencies of these classifications are recorded in the table below:

	only coffee	only tea	both	neither
Science	0.15	0.11	0.10	0.03
Business	0.12	0.09	0.04	0.02
Humanities	0.07	0.13	0.06	0.08

For the following questions, if you can read the answer directly off the table you do not need to show any work. If your answer uses a combination of values from the table you need to show your work. (For example, you could show your work by writing the values being used in a summation.)

- (a) What is the probability that a randomly selected student drinks only tea?

$$P(\text{only tea}) = 0.11 + 0.09 + 0.13 = 0.33$$

- (b) What is the probability that a randomly selected student is from the faculty of Business and drinks only coffee?

$$P(\text{Business} \cap \text{only coffee}) = 0.12$$

- (c) What is the probability that a randomly selected student drinks neither coffee nor tea or is from the faculty of Science?

$$P(\text{neither} \cup \text{Science}) = 0.03 + 0.02 + 0.08 + 0.15 + 0.11 + 0.10 = 0.49$$

or

$$P(\text{neither} \cup \text{Science}) = P(\text{neither}) + P(\text{Science}) - P(\text{neither} \cap \text{Science})$$

$$= (0.03 + 0.02 + 0.08) + (0.15 + 0.11 + 0.10 + 0.03) - 0.03 = 0.49$$

- (d) What is the probability that a randomly selected student is registered in Science or Humanities, and they drink only coffee?

$$P((\text{Science} \cup \text{Humanities}) \cap \text{only coffee})$$

$$= P(\text{Science} \cap \text{only coffee}) + P(\text{Humanities} \cap \text{only coffee})$$

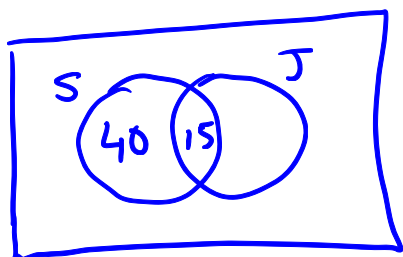
$$= 0.15 + 0.07$$

$$= 0.22$$

3. [2 marks] A high school graduating class consists of 120 students. Of these students, 110 of them are studying at least one of the three following languages: French, Spanish, and Japanese. In particular, the course enrollment is as follows: 60 are in French, 55 are in Spanish, 65 are in Japanese, 30 are in French and Spanish, 35 are in French and Japanese, 10 are in all three languages. You also know that there are 40 students who are taking Spanish but not Japanese.

- (a) Find the probability that a randomly selected student is enrolled in Spanish and Japanese. (Hint: a picture might be useful here!)

Say $F = \text{French}$, $S = \text{Spanish}$, $J = \text{Japanese}$



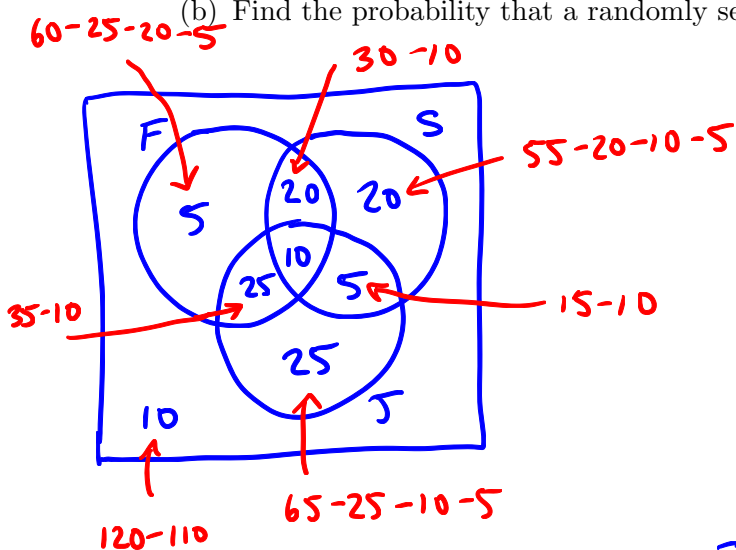
$$n(S) = 55, \quad n(J) = 65$$

$$n(S \cap \bar{J}) = 40$$

$$\begin{aligned} n(S \cap J) &= n(S) - n(S \cap \bar{J}) \\ &= 55 - 40 = 15 \end{aligned}$$

$$P(S \cap J) = 15/120 = 1/8 = 0.125$$

- (b) Find the probability that a randomly selected student is enrolled in exactly one language course.



$$P(\text{exactly one})$$

$$= P(F \cap \bar{S} \cap \bar{J}) + P(\bar{F} \cap S \cap \bar{J}) + P(\bar{F} \cap \bar{S} \cap J)$$

$$= \frac{5}{120} + \frac{20}{120} + \frac{25}{120}$$

$$= \frac{50}{120} = 5/12 = 0.4167$$