STAT 260 Summer 2024: Written Assignment 1

Due: Upload your solutions to Crowdmark BEFORE 6pm (PT) Friday May 17.

You may upload and change your files at any point up until the due date of Friday May 17 at 6pm (PT).

A 2% per hour late penalty will be automatically applied within Crowdmark. The penalty is applied in such a way so that assignments submitted 6pm to 6:59pm will have 2% deducted, assignments submitted 7pm-7:59pm will have 4% deducted, etc.

Note that if you submit any portion of your assignment before the deadline, Crowdmark will NOT permit you to edit your submission (including make additional uploads) after the 6pm deadline passes. This means that if, for example, you upload only Question 1 before the deadline, you will not be able to upload Question 2 after the deadline. If you intend to submit late (with penalty) you must submit the entire assignment late.

Submission: Solutions are to be uploaded to Crowdmark. Here you will be asked to upload your solutions to each question separately. Your solution to Question 1 must be uploaded in the location for Question 1, your solution to Question 2 must be uploaded in the location for Question 2, etc. If your work is uploaded to the wrong location, the marker will not be able to grade it.

You may hand-write your solution on a piece of paper or tablet. 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 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. If you complete your work on a tablet, save the file as a PDF or each question as a jpeg and drag/drop the file into the Crowdmark box. ***Photographs of laptop/tablet screens will not be graded***; take a proper screenshot.

Instructions: For full marks, your work must be neatly written, and contain enough detail that it is clear how you arrived at your solutions. ***You will be graded on correct notation***. Messy, unclear, or poorly formatted work may receive deductions, or may not be graded at all. Only resources presented in lecture or linked to on the Stat 260 Brightspace page are permitted for use in solving these assignments; using outside editors/tutors, and/or software (include AIs) is strictly forbidden. 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.

1. [**4 marks**] Suppose that the Labour Counsel for Oceanview City wants to determine the mean hourly wage for food and beverage servers in their municipality. They set up an online poll open to food and beverage servers working in any of the city's 290 licensed restaurants, and receive a total of 317 responses. From the survey, the lowest hourly wage was \$16.75 and the highest was \$34.75. The mean of the responses was \$19.75 with a standard deviation of \$1.20.

   Determine each of the following. Be very precise/specific in your solutions.

   (a) What is the population in this scenario?

   The population would be all food and beverage servers working in the 290 licensed restaurants.

   (b) What is the sample in this scenario?

   The sample comprises of the 317 servers that responded to the online poll.

   (c) What is the parameter of interest in this scenario? Include the correct notation (symbol) for the parameter.

   The mean hourly wage of all food and beverage servers in the city Oceanview, would be the parameter. Signified or Notation Symbol → μ (mu)

   (d) What is the random variable being measured? Is this a discrete or a continuous random variable?

   The random variable being measured is the hourly wage of food and beverage servers. It can be classified as a discrete random variable as wages have an infinitely listable amount of denotions or outcomes. However, money or currency at times can be considered as a continuous random variable to help ease calculations. That is not the case here though.

2. **[6 marks]** A sample of 8 Western Sandpiper (*Calidris mauri*) eggs are examined and found to have the following lengths in millimeters (mm):

318 319 309 317 295 313 316 308

(a) Find the sample mean $\bar{x}$ and sample median $\tilde{x}$. Show the full calculation, including correct notation.

Mean $\bar{x} = \dfrac{295+308+309+313+316+317+318+319}{8} = 311.875$

Median $\tilde{x} = \dfrac{313+316}{2} = 314.5$

~~295~~ ~~308~~ ~~309~~ 313 316 ~~317~~ ~~318~~ ~~319~~

$\dfrac{m_1+m_2}{2}$

(b) Find the sample standard deviation s. Show the full calculation, including correct notation.

$\bar{x}$ Mean found above $= 311.875$, $n = 8$

$\therefore s^2 = \left(\frac{1}{8-1}\right)\Big((295-311.875)^2+(308-311.875)^2+(309-311.875)^2+(313-311.875)^2+(316-311.875)^2$
$+(317-311.875)^2+(318-311.875)^2+(319-311.875)^2\Big)$

$= \left(\frac{1}{7}\right)(440.875) \approx 62.98$

Variance $(s^2) = 62.98$

standard deviation $(s) = \sqrt{62.98}$

$s = 7.936 \approx 7.94$

(c) Determine the proportion of the sample that lies within 2 standard deviations from the sample mean. That is, within the interval $[\bar{x} - 2s, \bar{x} + 2s]$.

$\bar{x} = 311.875 \quad s = 7.94$

$\bar{x} - 2s \rightarrow 311.875 - 2(7.94) = 296.0027 \approx 296$

$\bar{x} + 2s \rightarrow 311.875 + 2(7.94) = 327.7473 \approx 327.75$

Interval $= [296, 327.75]$

From the data 7/8 samples lie within the interval, 295 being the outlier.

$\therefore$ proportion $= \dfrac{\text{num of values within interval}}{\text{total number of values}} = \dfrac{7}{8} = 0.875$ or 87.5%

The proportion of the sample that lies within 2 standard deviations from the sample mean is 0.875 or 87.5%.

3. **[4 marks]** The concentration of total mercury (T-Hg) (in micrograms per gram ($\mu g/g$) of wet weight) was measured in the livers of four adult female killer whales (*Orcinus orca*). The smallest of the four observations was $x_1 = 38\ \mu g/g$, and the largest was $x_4 = 100\ \mu g/g$. The four observations yield a sample mean of $\bar{x} = 64\ \mu g/g$ and a sample variance of $s^2 = 680\ \mu g^2/g^2$.

Determine the values of the middle two observations x_2 and x_3 (assume $x_2 \leq x_3$). Show your work; do not just guess and check answers.

Hint: use the formulas for mean and variance to solve for the two unknowns (x_2 and x_3).

① $\bar{x} = \dfrac{38 + x_2 + x_3 + 100}{4} = 64$

$\rightarrow x_2 + x_3 + 138 = 256$

$\rightarrow x_2 + x_3 = 256 - 138$

$\rightarrow x_2 + x_3 = 118$

② $680 = \frac{1}{3}\left[(38-64)^2 + (x_2-64)^2 + (x_3-64)^2 + (100-64)^2\right]$

$2040 = 676 + (x_2-64)^2 + (x_3-64)^2 + 1296$

$\rightarrow (x_2-64)^2 + (x_3-64)^2 = 68$

∵ since $x_2 + x_3 = 118$
$x_3 = 118 - x_2$

Finding x_2: $(x_2-64)^2 + ((118-x_2)-64)^2 = 68$

$(x_2-64)^2 + (54-x_2)^2 = 68$

$(x_2)^2 - (128) + 4096 + (x_2)^2 - (108) + 2916 = 68$

$\rightarrow (x_2)^2 - (118) + 3472 = 0$

using $\dfrac{-b \pm \sqrt{b^2-4ac}}{2a}$ where

$a = 1$

$b = -118$

$c = 3472$

we get x_2 as $x_2 = 62$ or $x_2 = 56$

Using our intial assumption of $x_2 \geq x_3$ we can conclude, that $x_2 = 56$ and $x_3 = 62$ as our two middle observations