STAT 2290 **Exam 1B Solutions**, Feb 2026, 8 problems, 75 minutes

Read these instructions:

- Leaving the testing room results in a new exam given for unfinished problems.
- Three detached sheets of notes allowed.
- No electronics.
- Raise your hand for questions or more paper.

Page	**1**	**2**	**3**	**4**
Points	10	20	20	15
Score				

Problem 1.
(A) Compute the population standard deviation of the dataset $(x_1, x_2, x_3, x_4) = (6, 4, 4, 2)$.

Answer $\boxed{\sqrt{2}}$
The mean is $\dfrac{6+4+4+2}{4} = 4$. So the population variance is

$$\frac{1}{4}\|(6,4,4,2) - 4\|^2 = \frac{1}{4}\|(2,0,0,-2)\|^2 = \frac{1}{4}[4+0+0+4] = 2.$$

So the population standard deviation is $\sqrt{2}$.

(B) Compute the z-score of the data point $x_2 = 6$ in the above dataset.

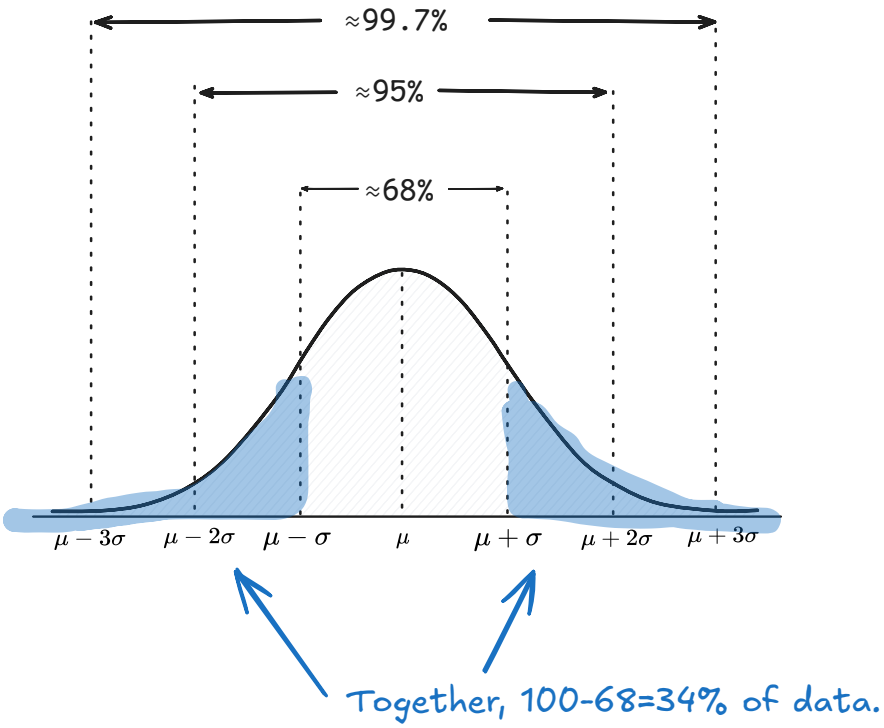
Answer $\boxed{\sqrt{2}}$
Correction: $x_2 = 6$ should read as $x_1 = 6$. The z-score of $x_1 = 6$ is

$$z = \frac{x_1 - \mu}{\sigma} = \frac{6-4}{\sqrt{2}} = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

Problem 2. Scores on the Wechsler Adult Intelligence Scale (a standard IQ test) for the 20 to 34 age group are approximately normally distributed with $\mu = 110$ and $\sigma = 25$.

(A) Roughly what percent of people aged 20 to 34 have IQs above 135?

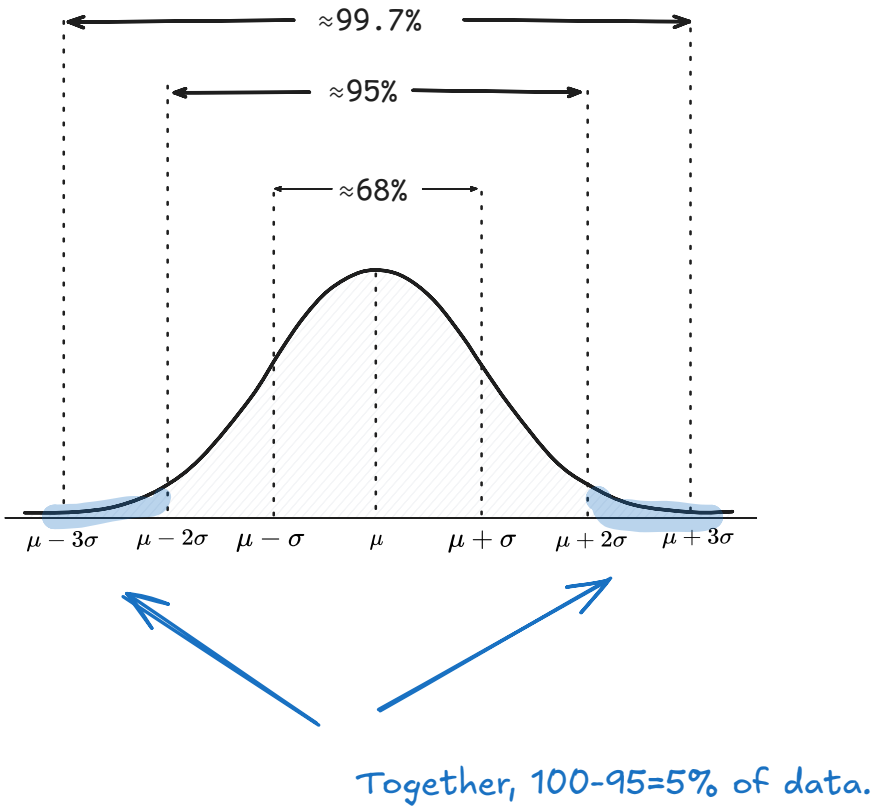
Answer: $\boxed{16\%}$
Since $\mu + \sigma = 110 + 25 = 135$, we know by the 68-95-99.7 rule that 68% of the data is positioned between $z = \mu - \sigma$ and $z = \mu + \sigma$ in the normal distribution. Therefore $100\% - 68\% = 32\%$ of the data lie outside of these bounds. So $32/2 = 16\%$ of data lie above $\mu + \sigma$.



(B) Roughly what percent of people aged 20 to 34 have IQs below 60?

Answer $\boxed{2.5\%}$

Since $\mu - 2\sigma = 110 - 2 \times 25 = 60$, we know by the 68-95-99.7 rule that 95% of the data is positioned between $z = \mu - 2\sigma$ and $z = \mu + 2\sigma$ in the normal distribution. Therefore $100\% - 95\% = 5\%$ of the data lie outside of these bounds. So $5/2 = 2.5\%$ of data lie below $\mu - 2\sigma$.



(C) Roughly what percent of people aged 20 to 34 have IQs between 60 and 135?

Answer $\boxed{83.97\%}$
The region under the normal distribution which is the complement of the regions of (A) and (B) is given by $100 - 16 - 0.03 = 83.97\%$.

Problem 3. Starbucks will study customer purchasing behavior to decide whether to expand its plant-based menu options nationwide. The analytics team proposes the following database queries to gather survey participants. State which sampling method each proposed query below corresponds to:

(A) Survey customers who respond to a promotional email about plant-based products during the first week of the campaign. $\boxed{\text{Voluntary response sampling}}$

(B) Randomly select 3,000 customers from the full national customer database: $\boxed{\text{SRS}}$

(C) Separate customers into two groups: Rewards members and non-members. Randomly select 1,500 customers from each group: $\boxed{\text{Stratified sampling}}$

(D) Randomly select 50 store locations across the country and survey every customer who makes a purchase at those selected stores on a particular day: $\boxed{\text{Cluster sampling}}$

Note: For (A), convenience sampling is also accepted. (A) and (B) only refer to the method of querying and not the surveying aspect and therefore will not permit other answers.

Problem 4. University of Helsinki researchers asked 305 healthy pregnant women to report their chocolate consumption. Six months after birth, the researchers asked mothers to rate their infants' temperament, including smiling, laughter, and fear. The babies born to women who had been eating chocolate daily during pregnancy were found to be more active and "positively reactive"—a measure that the investigators said encompasses traits like smiling and laughter.

(A) Was this an observational study or an experiment? Justify your answer.

Answer
This was an observation study: the pregnant women were only asked to report their chocolate consumption and infant temperament, and the researchers imposed no treatments on the subjects.

(B) Does this study show that eating chocolate regularly during pregnancy helps produce infants with good temperament? Explain briefly.

Answer
No: since this study was an observational study, it cannot establish or suggest causation.

Problem 5.
(A) Identify the sample space of the following probability experiment: Guessing the initial of a student's middle name:

Answer

Making the assumption such a middle name exists and that it is within the English alphabet, we can write the sample space as {A, B, C, D, E, ..., X, Y, Z}. Other reasonable answers are also accepted.

(B) Determine the size of the sample space in (A).

Answer

The sample space {A, B, C, D, E, ..., X, Y, Z} has size 26.
Reasonable answers that agree with your answer to part (A) are accepted.

Problem 6. Does drinking coffee make you more alert when studying for a test? To test this, an experiment will involve 50 volunteers who will take a memory test 20 minutes after drinking cola. Some volunteers will be randomly assigned to drink caffeine-free cola; some to drink regular cola (with caffeine), and the others a mixture of the two (getting a half dose of caffeine). For each volunteer, a test score (the number of items recalled correctly) will be recorded. The volunteers will not be told which type of cola they have been given, but the researcher who evaluates the results will prepare the cups of cola right on the spot (out of sight of the volunteers).

(A) Identify the treatments of this experiment. No explanations needed.

Answer

The three treatments are:

- caffeine-free cola,
- regular cola (with caffeine), and
- a mixture of the two (getting a half dose of caffeine).

(B) Identify the control group of this experiment, or explain why it does not exist here.

Answer

Subjects who are assigned to drink the caffeine-free cola.

(C) Is this experiment single-blinded? Briefly justify your answer. If it is not single-blinded, explain how to make it single-blinded.

Answer

Yes: The volunteer subjects are blinded from knowing the treatments given since "the researcher who evaluates the results will prepare the cups of cola right on the spot (out of sight of the volunteers)."

(D) State three ways to improve the above experiment. Justify each proposed improvement with a sentence.

Answer

Some sample answers include:

- Use a larger sample size than 50, since 50 subjects divided into 3 treatment groups results in roughly 17 subjects in each treatment group.
- Some subjects may recognize caffeinated vs. noncaffeinated colas. One alternate way to blind the subjects is to add other flavors to mask the caffeine taste and smell, or ask the subject what drink they think they just drank from a list of drink options.
- Make the study double-blinded by not letting the researchers who hand out the cola drinks to the subjects know which treatment is which. This reduces bias on the part of the experimenters.
- Gender or age may often act as a confounding variable that affects the results of such experiments; to counteract the effects of such confounding variables, we can **block** based on gender or age groups to strengthen our experimental results.
- Since we want to test the hypothesis on whether drinking caffeine helps with studying for a test, we should offer more caffeinated and noncaffeinated drink options than just colas (like tea, decaf. vs. caffeinated coffee). This reduces the chance that preference or non-preference for colas is a confounder in the study.
- Since subjects may naturally perform differently on memory tests, it is better to give each subject a memory test before the treatment and then another memory test after the treatment to compare how well they do before and after the treatment.

Problem 7. Write down numeric expressions that count the following tasks. Do not simplify factorials, binomial coefficients, sums, and products.

(A) How many ways are there to obtain exactly six tails by flipping a fair coin ten times? $\boxed{\binom{10}{6}}$

(B) How many ways can 1st and 2nd places be awarded to five runners in a race? (Assume there are no ties.) $\boxed{5 \times 4}$

(C) How many 4-digit numeric passcodes are composed of only even digits? Examples include "2422" and "0820". $\boxed{5 \times 5 \times 5 \times 5}$

(D) How many rearrangements of the word "MISSISSIPPI" are there? $\boxed{\frac{11!}{4!4!2!}}$

(E) How many 5 card hands drawn from a standard deck of 52 cards (evenly split into one of four suits: hearts, diamonds, clubs, spades) contain only cards of the heart suit? (Assume that the order in which you receive the cards does not matter: you only care about how many different 5-card possibilities there are.)

Answer

We choose 5 cards from the $52/4 = 13$ heart cards, or $\binom{13}{5}$.

Problem 8. A student randomly selected 60 boxes of different types of cereal from a large grocery store. One group of 31 cereal boxes showed servings of sugar in ¾-cups; its boxplot is displayed below. The remaining group of 29 cereal boxes showed servings of sugar in 1-cup: the data set is shown below:

```
00 02 02 02 03 03 04 04 04 07 08 12 12 13 13
13 14 14 14 14 15 15 15 16 16 17 18 19 20
```

Note

The 3/4-cup boxplots have been slightly modified after the exam, but this should not affect the answers.

(A) Draw into Figure 1 a boxplot of the above dataset of 29 cereals:

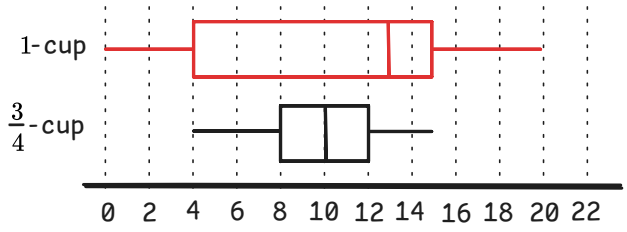


Figure 1. Sugar Content Per Serving

(B) Write a few sentences to compare the distributions of sugar content per serving for the 1-cup vs ¾-cups cereals.

Answer

Shape: While the 1-cup dataset is strongly left-skewed, the 3/4-cup dataset is roughly symmetric (with a very minor left skew).
Center: The median of the 1-cup dataset (13) is greater than that of the 3/4-cup dataset (≈ 10).
Spread: Cereal boxes with 1-cup labels have much greater variability (IQR =) in sugar content than cereal boxes with 3/4-cup labels (IQR≈ 4, Range≈ 11)
Outlier: Both datasets have no outliers.

To make a better comparison between the two groups of cereals, the student converted the ¾-cup servings data into 1-cup servings by multiplying by $4/3$. The bottom boxplot shown below displays sugar content (in grams) per cup for those cereals that showed a serving size of ¾-cup on their side panels:

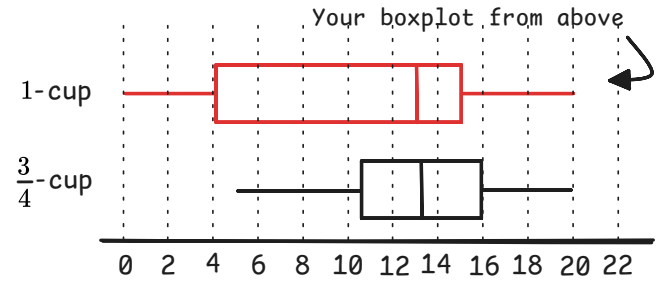


Figure 2. Adjusted Sugar Content Per Cup (grams)

(C) What new information about sugar content does Figure 2 provide?

Answer

The median sugar contents of both the 1-cup and 3/4-cup samples now appear to be roughly the same. However, cereal boxes with 1-cup labels still exhibit much greater variability in sugar content than boxes with 3/4-cup labels.

(D) Based on the boxplots shown above on this page, how would you expect the mean amounts of sugar per cup to compare for the different recommended serving sizes? Explain.

Answer

(During the exam, we clarified that this problem should be interpreted as follows: compare the means of the two

boxplots shown in Figure 2.)

The 3/4-cup boxplot is roughly symmetric, so its mean is roughly the same as the median of that boxplot (≈ 13). (You can argue that the boxplot is slightly left skewed, so the mean is actually slightly less than 13). The 1-cup boxplot exhibits a very strong left skew, so its mean is much less than the median of ≈ 13. In other words, the mean of the 1-cup boxplot is expected to be much less than the mean of the 3/4-cup boxplot.