Probability & Statistics Final Exam Solutions

krista king MATH

Probability & Statistics Final Exam Answer Key

1. (5 pts)	■	B	C	D	E
2. (5 pts)	A	■	C	D	E
3. (5 pts)	A	B	C	■	E
4. (5 pts)	A	B	■	D	E
5. (5 pts)	A	B	■	D	E
6. (5 pts)	A	B	C	■	E
7. (5 pts)	A	B	C	D	■
8. (5 pts)	■	B	C	D	E

9. (15 pts) Mean: $\mu = 79.69$, Median: 82, Mode: 88

10. (15 pts) $\sigma \approx 2.915$

11. (15 pts) $20.95\,\%$ chance of rain

12. (15 pts) $73.72\,\%$

Probability & Statistics Final Exam Solutions

1. A. The outlier of the data set is 6, because 6 is an extreme value compared to the rest of the data. Find the mean of the data set including the outlier.

$$\mu = \frac{6 + 31 + 40 + 42 + 44 + 53}{6} = \frac{216}{6} = 36$$

Find the mean of the data set after removing the outlier of 6.

$$\mu = \frac{31 + 40 + 42 + 44 + 53}{5} = \frac{210}{5} = 42$$

Find the difference between the means.

$$42 - 36 = 6$$

The mean increased by 6 when the outlier was removed.

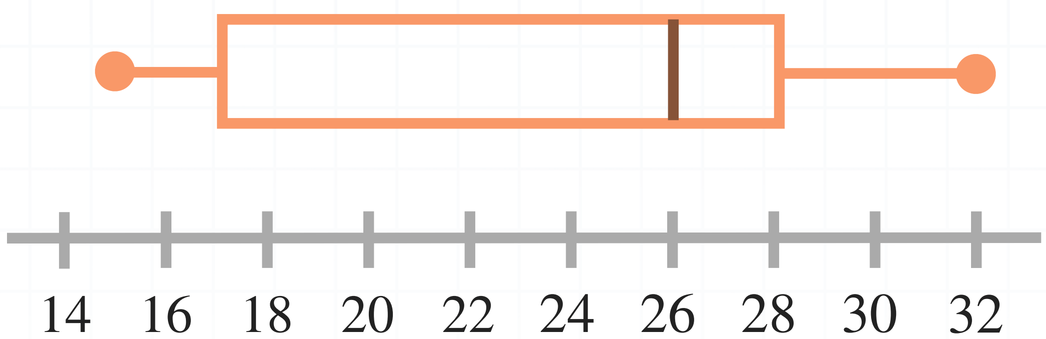
2. B. First organize the data from smallest to largest.

$$15,\ 17,\ 19,\ 26,\ 26,\ 28,\ 32$$

The minimum is 15 and the maximum is 32. These values will be the ends of the whiskers. The data is already in ascending order, so we can see that the median (the middle number) is 26. This will be the line in the middle of the box.

The lower half of the data is $15, 17, 19$, and the median of that lower half is $Q_1 = 17$. The upper half of the data is $26, 28, 32$, and the median of that upper half is $Q_3 = 28$.

Answer choice B has the correct box-and-whisker plot with the minimum at 15, Q_1 at 17, the median at 26, Q_3 at 28, and the maximum at 32.



3. D. We know that $\mu = 101.77$ and $\sigma = 0.41$. Since our competitor wants to finish in the top 3%, we'll look for the value in the body of the negative z-table that keeps us just under 0.0300. That value is 0.0294, which gives us a z-score of $z = -1.89$. Using the z-score formula, we then get

$$z = \frac{x - \mu}{\sigma}$$

$$-1.89 = \frac{x - 101.77}{0.41}$$

$$-0.7749 = x - 101.77$$

$$101.77 - 0.7749 = x$$

$$100.9951 = x$$

This means that the slowest possible time she can run is 100.995 seconds in order to finish among the top 3% of the runners.

4. C. Use the addition rule for probability.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A) = P(\text{jack}) = \frac{4}{52}$$

$$P(B) = P(\text{red}) = \frac{26}{52}$$

$$P(A \cap B) = P(\text{jack and red}) = \frac{2}{52}$$

Therefore, the probability that we get a jack or a red card is

$$P(A \cup B) = \frac{4}{52} + \frac{26}{52} - \frac{2}{52}$$

$$P(A \cup B) = \frac{28}{52}$$

$$P(\text{jack or red}) = \frac{7}{13}$$

5. C. This is a combination problem since the order in which we use the points doesn't matter. In this case, there are 5 points and we use 3 at a time to create a triangle.

$$_nC_r = \frac{n!}{r!(n-r)!}$$

$$_5C_3 = \frac{5!}{3!(5-3)!}$$

$$_5C_3 = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1(2!)}$$

$$_5C_3 = \frac{5 \cdot 4}{2!}$$

$$_5C_3 = \frac{20}{2}$$

$$_5C_3 = 10$$

6. D. This is a binomial random variable problem since 1) the marble is replaced and the events are independent, 2) pulling a blue marble is a success and pulling a yellow marble is failure, 3) there are a fixed number of trials: 10, and 4) the probability of success (pulling a blue marble) is constant on each trial.

To find the probability that exactly 7 blue marbles are pulled in 10 total pulls, use the formula:

$$P(r \text{ successes in } n \text{ attempts}) = \binom{n}{r} p^r(1-p)^{n-r}$$

$$P(7 \text{ blues in } 10 \text{ pulls}) = {_{10}C_7} \cdot \left(\frac{5}{7}\right)^7 \left(\frac{2}{7}\right)^3$$

We'll calculate the binomial coefficient.

$$_{10}C_7 = \frac{10!}{7!(10-7)!} = 120$$

Then the probability is

$$P(7 \text{ blues in } 10 \text{ pulls}) = 120\left(\frac{5}{7}\right)^7\left(\frac{2}{7}\right)^3$$

$$P(7 \text{ blues in } 10 \text{ pulls}) \approx 0.2655$$

$$P(7 \text{ blues in } 10 \text{ pulls}) \approx 27\,\%$$

7. E. Set up the inequality, including the finite population correction factor, since we have a finite population of $4{,}900$ bison in the herd. Since the margin of error is $\pm 5\,\%$, we'll put 0.05 on the right side.

$$z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}} \le 0.05$$

To find the smallest possible sample size n that keeps us within the margin of error, we need to optimize $\hat{p}(1-\hat{p})$ by making it as large as possible. This happens when $\hat{p} = 0.5$.

$$z^*\sqrt{\frac{0.5(1-0.5)}{n}} \le 0.05$$

For a $92\,\%$ confidence level, we're looking at the middle $92\,\%$ probability in a normal distribution. This means we'll have $4\,\%$

probability in each tail, and we're interested in the z-score that puts us at $100\% - 4\% = 96\%$ probability. Using a z-table, we find that 1.75 gets us a probability of $0.9599 \approx 0.9600$.

$$1.75\sqrt{\frac{0.5(1-0.5)}{n}}\sqrt{\frac{4{,}900-n}{4{,}900-1}} \le 0.05$$

Simplify the left side, then move everything but the n terms to the right.

$$1.75\sqrt{\frac{0.5(0.5)}{n}}\sqrt{\frac{4{,}900-n}{4{,}899}} \le 0.05$$

$$1.75\frac{\sqrt{(0.5)^2}}{\sqrt{n}}\frac{\sqrt{4{,}900-n}}{\sqrt{4{,}899}} \le 0.05$$

$$1.75\left(\frac{0.5}{\sqrt{4{,}899}}\right)\left(\frac{\sqrt{4{,}900-n}}{\sqrt{n}}\right) \le 0.05$$

$$\frac{\sqrt{4{,}900-n}}{\sqrt{n}} \le 0.05\left(\frac{\sqrt{4{,}899}}{0.5(1.75)}\right)$$

Simplify the left side, then square both sides to solve for n.

$$\sqrt{\frac{4{,}900-n}{n}} \le 0.05\left(\frac{\sqrt{4{,}899}}{0.5(1.75)}\right)$$

$$\sqrt{\frac{4{,}900}{n}-\frac{n}{n}} \le 0.05\left(\frac{\sqrt{4{,}899}}{0.5(1.75)}\right)$$

$$\sqrt{\frac{4,900}{n} - 1} \le 0.05\left(\frac{\sqrt{4,899}}{0.5(1.75)}\right)$$

$$\frac{4,900}{n} - 1 \le \left(\frac{0.05\sqrt{4,899}}{0.5(1.75)}\right)^2$$

$$\frac{4,900}{n} \le \left(\frac{0.05\sqrt{4,899}}{0.5(1.75)}\right)^2 + 1$$

$$\frac{4,900}{n} \le 16.9967$$

$$\frac{4,900}{16.9967} \le n$$

$$288.2907 \le n$$

$$n \ge 288.2907$$

Because we need to sample more than 288 of the bison, and we can't sample part of a bison, we'll need to sample 289 bison in order to stay within a $\pm 5\%$ margin of error at 92% confidence.

8. A. Find the slope and y-intercept using the correct formulas. It's helpful to create a chart of the values you'll need. We have 5 data points, so $n = 5$.

	x	y	xy	x^2
	0	0.8	0	0
	2	1.4	2.8	4
	4	3.1	12.4	16
	6	4.8	28.8	36
	8	6.2	49.6	64
Sum	**20**	**16.3**	**93.6**	**120**

Now we can plug everything into the formula for m,

$$m = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$m = \frac{5(93.6) - 20(16.3)}{5(120) - 20^2}$$

$$m = \frac{468 - 326}{600 - 400}$$

$$m = \frac{142}{200}$$

$$m = 0.71$$

and the formula for b.

$$b = \frac{(\sum y) - m(\sum x)}{n}$$

$$b = \frac{16.3 - 0.71(20)}{5}$$

$$b = \frac{2.1}{5}$$

$$b = 0.42$$

The line of best fit is given by the equation $y = 0.71x + 0.42$.

9. The measures of central tendency are mean, median, and mode. The mean is

$$\mu = \frac{\text{sum of all data points}}{\text{the number of data points}}$$

$$\mu = \frac{2{,}072}{26}$$

$$\mu = 79.69$$

If we line up all the data in ascending order, the two middle numbers are 81 and 83, so the median is the mean of these numbers.

$$\mu = \frac{81 + 83}{2} = 82$$

The mode is the number that occurs the most. In this case, the score 88 occurs 3 times, which is more than any other score.

10. To find standard deviation of a population, start by finding the mean.

$$\mu = \frac{4+7+9+12}{4} = \frac{32}{4} = 8$$

Find the difference between each data measure and the mean. Square the result.

$$4 - 8 = -4 \to (-4)^2 = 16$$

$$7 - 8 = -1 \to (-1)^2 = 1$$

$$9 - 8 = 1 \to 1^2 = 1$$

$$12 - 8 = 4 \to 4^2 = 16$$

Plug these results into the formula for standard deviation.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

$$\sigma = \sqrt{\frac{16+1+1+16}{4}} = \sqrt{\frac{34}{4}} = \sqrt{8.5}$$

$$\sigma \approx 2.915$$

11. Use Bayes' Theorem to find the probability that it will rain during your day at the beach.

$$P(A\,|\,B) = \frac{P(B\,|\,A) \cdot P(A)}{P(B)}$$

We're looking for the probability of rain given that it's already a cloudy morning, so $P(A\,|\,B) = P(\text{rain}\,|\,\text{cloud})$, and Bayes' theorem is

$$P(\text{rain}\,|\,\text{cloud}) = \frac{P(\text{cloud}\,|\,\text{rain}) \cdot P(\text{rain})}{P(\text{cloud})}$$

Each piece of Bayes' theorem is

$$P(\text{rain}) = 0.18$$

$$P(\text{cloud}) = 0.55$$

$$P(\text{cloud}\,|\,\text{rain}) = 0.64$$

So Bayes' theorem tells us that the probability that it will rain given that it's already a cloudy morning is

$$P(\text{rain}\,|\,\text{cloud}) = \frac{(0.64) \cdot (0.18)}{0.55}$$

$$P(\text{rain}\,|\,\text{cloud}) = \frac{0.1152}{0.55}$$

$$P(\text{rain}\,|\,\text{cloud}) = 0.209\overline{4545}$$

$$P(\text{rain}\,|\,\text{cloud}) \approx 20.95\,\%$$

12. First we need to verify normality. Remember that our sample space should be no more than $10\,\%$ of our population. The expected number of successes and failures should each be at least 10.

$$\frac{120}{5{,}000} = 0.024 = 2.4\,\%$$

This is less than $10\,\%$, so we've met the independence condition. We can also meet the normal condition:

$$120(0.6) = 72 \geq 10$$

$$120(0.4) = 48 \geq 10$$

The sample space was random so we've met the conditions of normality. Now we'll find the mean and standard deviation for the sample.

$$\mu_{\hat{p}} = p = 0.6$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.6(0.4)}{120}} \approx 0.0447$$

We need to find the probability that our results are within $5\,\%$ of the population proportion $p = 60\,\%$. This means, how likely is it that the mean of the sample proportion falls between $55\,\%$ and $65\,\%$? We can express $5\,\%$ in terms of standard deviations.

$$\frac{0.05}{0.0447} \approx 1.12$$

This means we want to know the probability of $P(-1.12 < z < 1.12)$. Using a z-table, -1.12 gives us 0.1314 and 1.12 gives us 0.8686.

$$P(-1.12 < z < 1.12) = 0.8686 - 0.1314$$

$$P(-1.12 < z < 1.12) = 0.7372$$

There's a 73.72% chance that our sample proportion will fall within 5% of the pizza shop's claim.

W W W . K R I S T A K I N G M A T H . C O M