

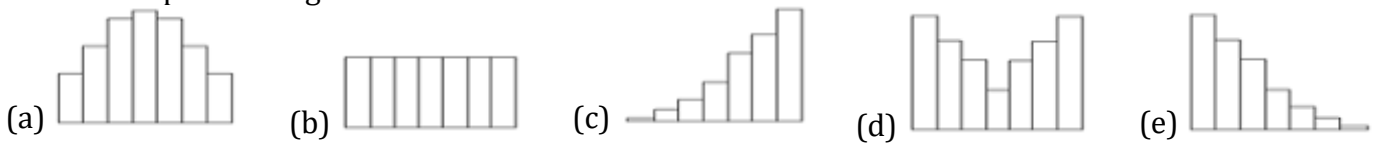
# MDM4U Review Questions Day 2

Put your name on the scantron. Do it now.

Use a pencil. Erase fully. Bubbling and erasing errors are your mistakes, not the machine's.

- If your experiment does not have a control group, which bias is present in your study?  
 (a) Hawthorne Effect (b) Implied Causation (c) Experimenter Mistakes (d) Blinding (e) Placebo Effect
- To find subjects for your MDM4U final project, you randomly select a time and head to the cafeteria to ask some questions. Which bias is present in your study?  
 (a) Transferring Findings (b) Convenience Sampling (c) Social Desirability Bias (d) Placebo Effect (e) Double Blinding
- You use a computer to randomly select 12 people from the phone book. A friend flips a coin to select which group each of your subjects will be assigned to. They record who is in each group and you don't look at it, so you are impartial during the study. Which bias applies to your study?  
 (a) Lack of Blinding (b) Placebo Effect (c) Lack of Double Blinding (d) Low Replication (e) Convenience Sampling
- Identify the type of graph you should use when you are doing two-variable, causal analysis.  
 (a) Histogram (b) Box and Whisker (c) Scatterplot (d) Pie Chart (e) Venn Diagram
- Identify the type of graph you should use when creating a probability distribution.  
 (a) Histogram (b) Box and Whisker (c) Scatterplot (d) Pie Chart (e) Venn Diagram
- Which is a measure of spread?  
 (a) Mean (b) Histogram (c) Standard Deviation (d) Z-Score (e) Expected Value

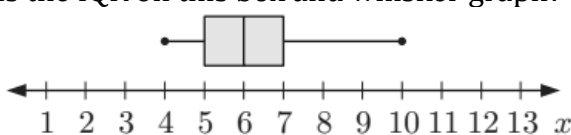
7. Which shape of histogram has mean < median < mode?



8. Which shape of histogram has mean = median = mode?



9. What is the IQR on this box and whisker graph?



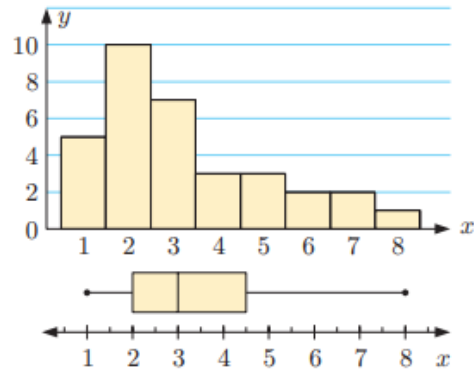
- (a) 1 (b) 4 (c) 6 (d) 10 (e) none of the previous

10. Q1 (quartile 1) is which percentile?

- (a) 0 (b) 0.25 (c) 0.4 (d) 4<sup>th</sup> (e) 25<sup>th</sup>

A statistician created this histogram and box and whisker diagram from the same set of data.

She was measuring the number of servings of vegetables that each of her subjects ate on Dec 14, 2012.



11. What is the Q1 on the above box and whisker?

- (a) 1                      (b) 2                      (c) 3                      (d) 4.5                      (e) 8

12. What is the mode on the above diagrams?

- (a) 1                      (b) 2                      (c) 3                      (d) 4.5                      (e) 8

13. What is the median on the above diagrams?

- (a) 1                      (b) 2                      (c) 3                      (d) 4.5                      (e) 8

14. What is the Q3 on the above box and whisker diagram?

- (a) 1                      (b) 2                      (c) 3                      (d) 4.5                      (e) 8

	A	B	C	D	E	F	G	H	I
1	Height	Age			Mean	Median	Mode	Minimum	Maximum
2	1.3	3		Height	2.0333	2	#N/A	1.3	3.1
3	1.5	5		Age	8.5	8	8	3	15
4	1.9	8							
5	2.1	8		X	Y	r	r <sup>2</sup>	slope	y-intercept
6	2.3	12		Height	Age	0.97532	0.9512	6.720779	-5.1655844
7	3.1	15		Age	Height	0.97532	0.9512	0.141538	0.83025641

15. In the above spreadsheet, what formula would appear in the cell E2?

- (a) =median(A2:A7)                      (d) =avg(B2:B7)  
 (b) =mean(A2:A7)                      (e) None of the previous.  
 (c) =avg(A2:A7)

16. In the above spreadsheet, what formula would appear in the cell H2?

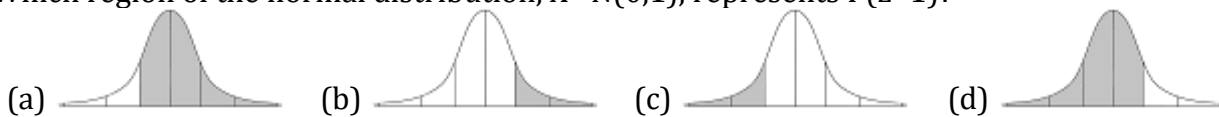
- (a) =min(A2:A7)                      (d) =smallest (A2:A7)  
 (b) =m(A2:A7)                      (e) None of the previous.  
 (c) =minimum(A2:A7)

17. In the above spreadsheet, what formula should appear in the cell F6?

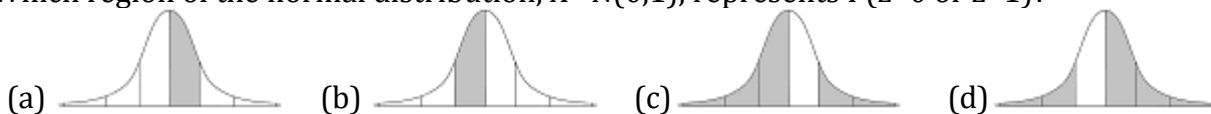
- (a) =r(A2:A7, B2:B7)                      (d) =correl(B2:B7, A2:A7)  
 (b) =r(B2:B7, A2:A7)                      (e) =correlcoeff(A2:A7, B2:B7)  
 (c) =correl(A2:A7, B2:B7)

18. Which is not a characteristic of the normal distribution?
- (a) It models the frequency and probability distributions of continuous random variables.
  - (b) It can be used to model a geometric probability distribution.
  - (c) The curve is symmetrical about the mean.
  - (d) The mean and median are equal.
  - (e) Approximately 68% for the data values are within one standard deviation.
19. On a normal distribution, what percentage of data falls within two standard deviations of the mean?
- (a) 13.5%
  - (b) 34%
  - (c) 95%
  - (d) 97%
  - (e) 99.7%
20. Assuming a normal distribution, what is  $P(z < 1.73)$ ?
- (a) 0.9554
  - (b) 0.9573
  - (c) 0.9582
  - (d) 0.9591
  - (e) 0.9664
21. Assuming a normal distribution, what is  $P(z > -1.45)$ ?
- (a) 0.0735
  - (b) 0.0156
  - (c) 0.2983
  - (d) 0.3498
  - (e) 0.9265
22. Assuming a normal distribution, what is  $P(z > -1 \text{ and } z < 0)$ ?
- (a) 0.1587
  - (b) 0.3413
  - (c) 0.500
  - (d) 0.6587
  - (e) 0.8413
23. The weight of adult male grizzly bears living in the wild in North America is approximately normally distributed with a mean of 220 kg and a standard deviation of 22 kg. How many standard deviations above the mean is a bear who weighs 250 kg?
- (a) 0
  - (b) 1
  - (c) 1.2
  - (d) 1.23
  - (e) 1.36
24. Gina's doctor told her that the z-score for her blood pressure, as compared to the blood pressure of other women her age, is 1.50. Which of the following is the best interpretation of this standardized score?
- (a) Gina's blood pressure is 1.50.
  - (b) Gina's blood pressure is 1.50 above the average blood pressure of women her age.
  - (c) Gina's blood pressure is 1.50 times the average blood pressure for women her age.
  - (d) Gina's blood pressure is 1.50 standard deviations above the average blood pressure of women her age.
  - (e) Only 1.5% of women Gina's age have a higher blood pressure than she does.

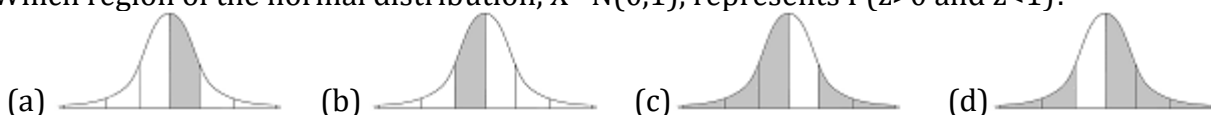
25. Which region of the normal distribution,  $X \sim N(0,1)$ , represents  $P(z > 1)$ ?



26. Which region of the normal distribution,  $X \sim N(0,1)$ , represents  $P(z < 0 \text{ or } z > 1)$ ?



27. Which region of the normal distribution,  $X \sim N(0,1)$ , represents  $P(z > 0 \text{ and } z < 1)$ ?



The next few questions consider the formulas in this spreadsheet:

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	The Data:					The Analysis:								
2	6	9	10	3		Mean	8.5	Minimum	3	Q3	10	X	5	
3	8	10	11	10		Standard Deviation	2.06	Q1	8	Maximum	11	Zscore	-1.7	
4	9	9	9	8		Mode	9	Median	9	IQR	2	P(x<5)	0.0447	

28. In the above spreadsheet, what formula should appear in the cell G3?

- (a) =stdev.p(A2:D4)
- (b) =stdv(A2:D4)
- (c) =σ(A2:D4)
- (d) =sd(A2:D4)
- (e) =standarddeviation(A2:D4)

29. In the above spreadsheet, what formula should appear in the cell I3?

- (a) =quartile(A2:D4)
- (b) =quartile1(A2:D4)
- (c) =quartileone(A2:D4)
- (d) =Q1(A2:D4)
- (e) =quartile.exc(A2:D4, 1)

30. In the above spreadsheet, what formula should appear in the cell M3?

- (a) =(M2-G2)/G3
- (b) =zscore(M2, G2, G3)
- (c) =zscore(G3, G2, M2, true)
- (d) =zscore(G3, G2, M2)
- (e) =zscore(G2, G3, M2)

31. In the above spreadsheet, what formula should appear in the cell M4?

- (a) =norm.dist(true, M2, G2, G3)
- (b) =norm.dist(true, \$M\$2, G2, G3)
- (c) =norm.dist(true, \$M\$2, \$G\$2, G3)
- (d) =norm.dist(M2, G2, G3, true)
- (e) =norm(true, M2, G2, G3)

32. How many ways can you spell the word snow travelling from top to bottom?



- (a) 2
- (b) 4
- (c) 8
- (d) 16
- (e) None of the previous

33. How many ways can you spell the word snows travelling from top to bottom?



- (a) 2
- (b) 4
- (c) 8
- (d) 16
- (e) None of the previous

34. Identify the distribution that represents the number of failures before you get a success in a series of Bernoulli trials.  
 (a) Normal (b) Binomial (c) Geometric (d) Hypergeometric (e) Uniform
35. Identify the distribution that gives the discrete probability distribution of obtaining exactly  $x$  successes out of  $n$  Bernoulli trials.  
 (a) Normal (b) Binomial (c) Geometric (d) Hypergeometric (e) Uniform
36. Identify the distribution that is a very common continuous probability distribution often used in the natural and social sciences to represent real-valued random variables.  
 (a) Normal (b) Binomial (c) Geometric (d) Hypergeometric (e) Uniform
37. Identify the distribution that has constant probability.  
 (a) Normal (b) Binomial (c) Geometric (d) Hypergeometric (e) Uniform
38. Identify the distribution that is a discrete probability distribution that describes the probability of successes in draws, without replacement.  
 (a) Normal (b) Binomial (c) Geometric (d) Hypergeometric (e) Uniform

This spreadsheet was created to simulate the number of times a person has to roll a dice until they get doubles. Recall that  $P(\text{doubles}) = 1/6$ .

	A	B	C	D	E	F	G	H	I
1	X	0	1	2	3	4	5	6	7
2	$q^x$	1	0.83333	0.69444	0.5787	0.48225	0.40188	0.3349	0.27908
3	$p$	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667
4	$P(x)=q^x \cdot p$	0.16667	0.13889	0.11574	0.09645	0.08038	0.06698	0.05582	0.04651

39. In the above spreadsheet, what formula should appear in the cell B2?  
 (a)  $=(5/6)^0$  (d)  $=(5/6)^{B1}$   
 (b)  $=q^x$  (e)  $=$A$1^{B1}$   
 (c)  $=1$
40. In the above spreadsheet, what formula should appear in the cell B3?  
 (a)  $=1-B2$  (d)  $=1/6$   
 (b)  $=0.16667$  (e)  $=$A$1$   
 (c)  $=\text{doubles}()$
41. In the above spreadsheet, what formula should appear in the cell B4?  
 (a)  $=B2 \cdot B3$  (d)  $=$B$2 \cdot $B$3$   
 (b)  $=$B$2 \cdot B3$  (e)  $=B2 \cdot B3$   
 (c)  $=B2 \cdot $B$3$

42. Which formula finds the probability of a single event in a geometric distribution?  
 (a)  $P(x) = C(n, x) \times p^x \times q^{(n-x)}$  (c)  $P(x) = q^x p$  (e)  $P(x) = n(q) \times n(p)$   
 (b)  $P(x) = \frac{C(a, x) \times C(n-a, r-x)}{C(n, r)}$  (d)  $P(x) = \frac{1}{n}$

43. From a carton of 24 eggs, of which 14 are cracked, 3 eggs roll out of the carton at random. What formula would be used to find the probability that exactly 2 of the rolling eggs are cracked?

- (a)  $P(x) = C(n, x) \times p^x \times q^{(n-x)}$       (c)  $P(x) = q^x p$       (e)  $P(x) = n(q) \times n(p)$   
 (b)  $P(x) = \frac{C(a, x) \times C(n-a, r-x)}{C(n, r)}$       (d)  $P(x) = \frac{1}{n}$

44. Which formula finds the probability of a single event in a binomial distribution?

- (a)  $P(x) = C(n, x) \times p^x \times q^{(n-x)}$       (c)  $P(x) = q^x p$       (e)  $P(x) = n(q) \times n(p)$   
 (b)  $P(x) = \frac{C(a, x) \times C(n-a, r-x)}{C(n, r)}$       (d)  $P(x) = \frac{1}{n}$

45. Which description of  $x$  is for a discrete random variable which could be used to make a binomial distribution?

- (a)  $x$  is a person's IQ score.  
 (b)  $x$  is the probability that the life of a lightbulb is between 90 and 115 hours.  
 (c) You draw  $x$  cards from a deck (without replacement) until you get a heart.  
 (d)  $x\%$  of people's heights are greater than 1.2 m.  
 (e) Guessing  $x$  true-false questions correctly on an examination with 5 questions.

46. A child flips a coin 5 times and counts the number of heads. Which formula finds the expected number of heads in the 5 coin flips?

- (a)  $E(x) = \sum P(x) \times x$       (c)  $E(x) = \frac{q}{p}$       (e)  $E(x) = \bar{x} = n \times p$   
 (b)  $E(x) = \frac{\min(x) + \max(x)}{2}$       (d)  $E(x) = \frac{ra}{n}$

47. A minivan has 6 seats. There are 18 people at a family picnic, 8 adults and 10 children. Six people are selected at random to go for ice cream. Which formula would be used to find the expected number of adults in the van?

- (a)  $E(x) = \sum P(x) \times x$       (c)  $E(x) = \frac{q}{p}$       (e)  $E(x) = \bar{x} = n \times p$   
 (b)  $E(x) = \frac{\min(x) + \max(x)}{2}$       (d)  $E(x) = \frac{ra}{n}$

48. Which formula would find the expected value in the winnings at a lottery?

- (a)  $E(x) = \sum P(x) \times x$       (c)  $E(x) = \frac{q}{p}$       (e)  $E(x) = \bar{x} = n \times p$   
 (b)  $E(x) = \frac{\min(x) + \max(x)}{2}$       (d)  $E(x) = \frac{ra}{n}$

49. A pair of dice is rolled until the player gets doubles. Which formula finds the expected number of times a player must roll the dice before they get doubles?

- (a)  $E(x) = \sum P(x) \times x$       (c)  $E(x) = \frac{q}{p}$       (e)  $E(x) = \bar{x} = n \times p$   
 (b)  $E(x) = \frac{\min(x) + \max(x)}{2}$       (d)  $E(x) = \frac{ra}{n}$

50. You roll an 8 sided die in an exciting game of dungeons and dragons. What is the probability of rolling a 1?  
 (a) 0.125      (b) 1/6      (c) 0.15625      (d) 0.53      (e) 0.8
51. There are 6 clear marbles and 4 striped ones in a bag. The distribution measures the number of striped marbles when 2 marbles are chosen without replacement. What is the probability of getting exactly one striped marble?  
 (a) 0.125      (b) 1/6      (c) 0.15625      (d) 0.53      (e) 0.8
52. There are 6 clear marbles and 4 striped ones in a bag. The distribution measures the number of striped marbles when 2 marbles are chosen without replacement. What is the expected number of striped marbles?  
 (a) 0.125      (b) 1/6      (c) 0.15625      (d) 0.53      (e) 0.8
53. A child flips a coin 5 times and counts the number of heads. Which formula would find the standard deviations for this probability distribution?  
 (a)  $\sigma = \frac{\sum f \times x}{\sum f}$       (c)  $\sigma = \sqrt{npq}$       (e) None of the previous would work for this distribution.  
 (b)  $\sigma = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}$       (d)  $\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$
54. In a sample of 200 tax returns, 40 contain errors. A government auditor selects a random sample of 3 (without replacement) to study. Which formula would find the standard deviation in this probability distribution?  
 (a)  $\sigma = \frac{\sum f \times x}{\sum f}$       (c)  $\sigma = \sqrt{npq}$       (e) None of the previous would work for this distribution.  
 (b)  $\sigma = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}$       (d)  $\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$
55. Circuit boards are assembled by selecting 4 computer chips at random from a large batch of chips. In this batch of chips, 90% of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 4 chips from this batch. What is the least probable value of X?  
 (a) 0      (b) 1      (c) 2      (d) 3      (e) 4
56. A non-profit organization plans to hold a raffle to raise funds for its operations. A total of 1,000 raffle tickets will be sold for \$1.00 each. After all the tickets are sold, one ticket will be selected at random and its owner will receive \$50.00. The expected value for the net gain for each ticket is -\$0.95. What is the meaning of the expected value in this context?  
 (a) The ticket owners lose an average of \$0.05 per raffle ticket.  
 (b) The ticket owners lose an average of \$0.95 per raffle ticket.  
 (c) Each ticket owner will lose exactly \$0.95 per raffle ticket.  
 (d) A ticket owner would have to purchase 19 more tickets for the expected value of his or her net gain to increase to \$0.00.  
 (e) A ticket owner has a 95 percent chance of having a ticket that is not selected.





64. A random sample of 432 voters revealed that 100 are in favour of a new law changing the colour of stop signs. A 95 percent confidence interval ( $z = 1.960$ ) for the proportion of the population of voters who are in favour of the new stop signs is:

(A)  $100 \pm 1.96\sqrt{\frac{0.5(0.5)}{432}}$       (C)  $100 \pm 1.96\sqrt{\frac{0.231(0.769)}{432}}$       (E)  $0.231 \pm 1.645\sqrt{\frac{0.231(0.769)}{432}}$

(B)  $100 \pm 1.645\sqrt{\frac{0.5(0.5)}{432}}$       (D)  $0.231 \pm 1.96\sqrt{\frac{0.231(0.769)}{432}}$

65. A market research firm asked 300 people who they were going to vote for in the upcoming election. Fifty-five people said they would vote for the rhinoceros party. Determine a 95% ( $z = 1.960$ ) confidence interval for the percent of people who will vote for the rhinoceros party.

- (a) 0                      (b) 0.0438                      (c) 0.183                      (d) 14 to 23%                      (e)  $18\% \pm 2\%$