MacEwan University

Stat 161

Midterm sample questions

1. The following table gives the first lines of a data set summarizing orders placed by a catering company.

food	amount ordered	number of boxes	quality (rating 1-5)
bananas	500 kg	50	4
apples	250 kg	10	5
pine apples	100 kg	20	2
bananas	200 kg	19	5
⋮	⋮	⋮	⋮

   (a) For each variable in the data set indicate if it is categorical, numerical discrete, or numerical continuous.

   (b) For each variable indicate which parts should be included with a complete data description.

2. Sketch a Venn diagram showing two events A and B. Show "A or not$B$" in the diagram.

3. Give an example for a discrete distribution and an example for a continuous distribution.

4. Sketch a bimodal, left skewed histogram.

5. How can you use the 5-number summary for a data set to locate the middle 50% and the top 50% in a sample?

6. What does it mean that the median is robust towards outliers?

7. What is described by the central limit theorem? What does the central limit theorem tell us about the shape?

8. What is the difference between a parameter and a statistic?

9. What is the purpose of Inferential Statistics?

10. What can you conclude about the shape of a histogram if the mean is much larger than the median for that sample?

11. Here are the IQ test scores of 29 seventh grade girls from an Albertan elementary school:

72	74	86	89	91	96	98	102	103	103
103	104	105	107	108	110	111	111	112	112
114	114	114	119	119	120	128	130	132	

   (a) Use an appropriate graph to visualize the data. Comment on your result.

   (b) Give the 5-number summary, and draw a box-plot.

(c) Find the mean and the standard deviation of the IQ scores. Use that $\sum x_i = 3087$, and $\sum x_i^2 = 334511$.

12. Don Iveson received 61.9% of the votes in the election for mayor of Edmonton on October 21, 2013.

    Assume you randomly choose 10 voters.

    (a) What is the probability that none of the 10 voted for Iveson?

    (b) What is the probability that at least one voted for Iveson?

13. How often do you floss your teeth per day?
    In a survey 55% of undergraduate students said they never floss, 32% floss once per day, 10% floss twice per day, and everyone else flosses three times per day.

    (a) Give the distribution of $X=$ number of times undergraduate students floss their teeth per day.

    (b) If an undergraduate is chosen at random, what is $P(X < 2)$?

14. The mean movie playtime is $\mu = 112$ minutes with a standard deviation of $\sigma = 18$ minutes. Assume that movie play times are normally distributed.

    (a) What is the probability that a randomly chosen movie is shorter than 102 minutes?

    (b) What is the probability that a randomly chosen movie is exactly 110 minutes long?

    (c) According to this model how long are the longest 30% of movies?

    (d) What is the probability that the AVERAGE playtime of 9 movies is shorter than 102 minutes.

15. There are 7 pizzas in an oven: 3 of the 7 pizzas have thick crust, of these one has only cheese and the other two have only mushrooms. The remaining 4 pizzas have a regular crust, and of these two have only cheese, and two have only mushrooms.

    Choose a pizza at random from the oven.

    (a) Find the probability for {getting a pizza with mushrooms}.

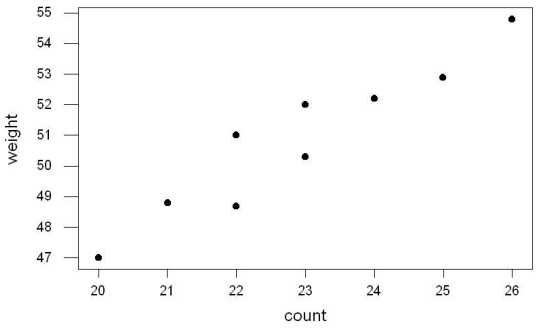
    (b) Find the probability for {getting a pizza with mushrooms & a thick crust }.

    (c) Find the probability for {getting a pizza with mushrooms | a thick crust }.

    (d) Are the events {getting a pizza with mushrooms} and {getting a thick crust} independent?

16. In order to study the relationship between count and weight of candies someone bought 9 bags chocolate coated peanuts and determined the number of candies and the weight of the candies included in each bag.

    The measurements are displayed in the following scatterplot



Summary statistics:(x = count, y = weight)

$$\sum x_i = 206 \quad \sum x_i^2 = 4744 \quad \sum x_i y_i = 10511$$
$$\sum y_i = 457.7 \quad \sum y_i^2 = 23324 \quad n = 9$$
$$s_x = 1.9 \quad s_y = 2.44$$

(a) Comment on the relationship between the count and the weight of the candies in a bag.

(b) Find Pearson's Correlation Coefficient. What can be learned from the result?

(c) Find the regression line describing the relationship between count and weight of the candies per bag.

(d) Interpret the intercept and the slope of the regression line.

(e) Would it be appropriate to use the regression line to estimate the weight of a bag of candies for a count of 25? Explain.

17. The following table gives the results from a survey. The table shows Age group ($20-29$, $30-39$, $40-49$) and how much an individual enjoys classical music (no, somewhat, yes).

	Enjoys Classical Music		
Age group	no	somewhat	yes
$20-29$	20	10	3
$30-39$	15	11	8
$40-49$	12	11	10

If an individual is randomly chosen from the table:

(a) Find the probability that the individual is in age group 20-29 AND enjoys classical music (yes).

(b) Find the probability that the individual is in age group 20-29.

(c) Find the probability that the individual enjoys classical music given he/she is in age group 20-29.

(d) Are the events being in age group 20-29 and enjoying classical music independent?

18. A study on tipping behaviour in a city showed that in the 30 sampled restaurants the average tip was 13.5% of the bill, the standard deviation is known to be 3%.

    (a) Give a 95% confidence interval for the mean tipping percentage in that city.

    (b) What is the margin of error?

    (c) Interpret your results from (a) and (b).