Answer Key

Math 30 Applied Ch 2 Statistics – Review Problems – Government Exam Style:

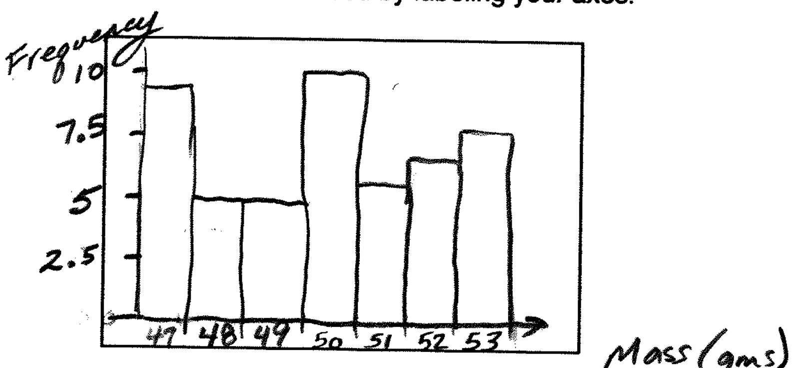
1. Skittles are packaged in bags that are labeled 50g. Fifty bags from the vending machine were tested. The actual amount of skittles in each bag was accurately measured and recorded in the table below.

- Calculate the probability for each mass and record these values in the chart.

Mass in grams (L_1)	Frequency (L_2)	Probability
47	9	9/50 = 0.18
48	5	5/50 = 0.10
49	5	5/50 = 0.10
50	10	10/50 = 0.20
51	6	6/50 = 0.12
52	7	7/50 = 0.14
53	8 /50	8/50 = 0.16

- When you add up the probabilities, they equal __1__.

- Construct a histogram to show the **probability** distribution for this experiment. Sketch the histogram and indicate the window used by labeling your axes.



- Calculate the mean and standard deviation of this data.

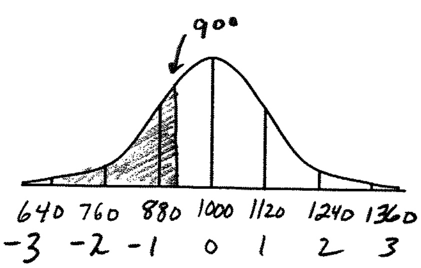
use 1 var-stats L_1, L_2 $\mu = 50.04$ $\sigma = 2.06$

- Explain what might happen to the company if the mean was too high or too low.

too low - rip off customers + they won't buy your skittles
too high - loose profits.

2. The lifetime of the X-Box 360 is normally distributed with a mean of 1000 hours and a standard deviation of 120 hours.

- Complete the normal curve below to show the distribution of the lifetime of the X-Boxes. Mark the points that are 1, 2, and 3 standard deviations form the mean.



- On your sketch, shade the area that represents choosing an X-box with a lifetime less than 900 hours.

- What is the probability of choosing at random an X-Box 360 with a lifetime less than 900 hours?

Normalcdf (0, 900, 1000, 120) = 0.20

- Approximately how many X-Boxes in a random sample of 10 000 would have a lifetime less than 700 hours?

Normalcdf (0, 700, 1000, 120) = 0.00620 × 10000 = 62

Use the following information to answer the next question

In humans, a particular pair of genes genetically determines whether your belly button is an innie or an outie. A person may have two dominant genes (BB), two recessive genes (bb) or a dominant and a recessive gene (Bb). If a person has one or two dominant genes then he or she will have an innie.

3. Complete the chart below to show the sample space for an offspring of parents who each carry one dominant (E) gene and one recessive gene.

		Mother	
		B	b
Father	B	BB	Bb
	b	Bb	bb

- What is the probability that an offspring from these parents will have an outie belly button?

$P(\text{outie}) = \frac{1}{4}$

Use the following information to answer the next part of the question

Approximately 16% of **all** people have an outie belly button.

- Calculate the mean and standard deviation for the number of people in a sample of 8700 that you would expect to have an outie belly button. Round your answer to the nearest hundredth.

$\mu = 0.16(8700) = 1392$ $\qquad \sigma = \sqrt{8700(0.16)(0.84)}$

$\sigma = 34.19$

- Calculate the symmetric 95% confidence interval for the number of people in this sample that have an outie belly button.

Lower

$\mu - 1.96(\sigma)$

$1392 - 1.96(34.19)$

1324.99

Upper

$\mu + 1.96\sigma$

$1392 + 1.96(34.19)$

1459.01

With 95% confidence I would expect between 1324 and 1460 people would have an outie bellie button.

Reviewing Binomial and Uniform Probability Distributions

1. The minimum daily temperatures in Edmonton during an 12 day period are

8.6, 7.3, 10.7, 15.2, 1.0, 5.9, 9.3, 8.6, 7.3, 8.5, 7.3, 7.5

Calculate, to the nearest tenth, the mean, median, mode, range, and standard deviation.

sto → L_1 use 1-var stats. mean = 8.1 mode = 7.3

med = 8 range = 15.2 − 1 = 14.2

θ = 3.12

2. The monthly sales of poutine at the STA cafeteria during a period of 28 days is listed.

380	350	430	270	240	400	380
210	340	270	350	360	240	330
270	330	440	450	230	330	330
240	330	270	320	350	290	310

Calculate the mean and standard deviation, respectively, of the data above.

a) Organize the data from lowest to highest in a straight line from left to right.

210 230 240 240 240 / 270 270 270 270 290 310 320 330 330 330 330 330 340 350 350 350 360 380 380. / 400 430 440 450

b) Calculate the percent of sales, to the nearest whole number, that are:

i) within one standard deviation of the mean, (i.e. in the interval $\mu \pm 1\theta$)

322.86 ± 63.35

$259.5 - 386.72$ $\frac{19}{28} = 0.68 = 68\%$

ii) within two standard deviations of the mean, (i.e. in the interval $\mu \pm 2\theta$).

$\mu - 2(63.35)$ $\mu + 2(63.35)$

$322.86 - 2(63.35)$ $322.86 + 2(63.35)$ $\frac{27}{28}$

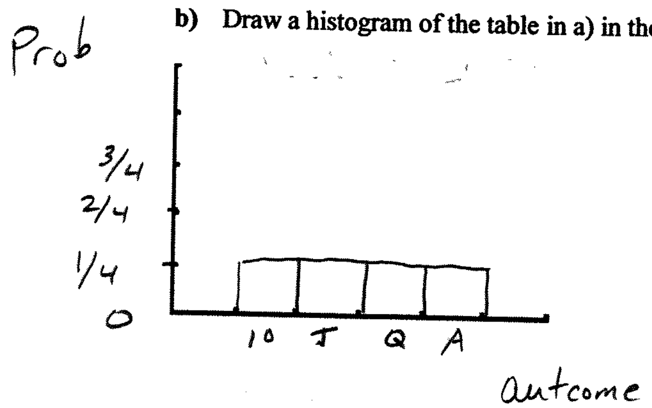
196.16 — 449.56 $= 96\%$

3. Four playing cards are placed in a pile. The pile consists of a ten of hearts, a jack of spades, a queen of diamonds and an ace of hearts. Consider the experiment a card is drawn and then the results recorded.

a) Illustrate all the possible outcomes and their respective probabilities by completing the table.

Outcome	Probability
10 H	0.25
J S	0.25
Q D	0.25
A H	0.25

b) Draw a histogram of the table in a) in the space below: (Be sure to label your axes.)



4. Lian Nails Ltd. produces 3-inch galvanized nails in boxes of 2000 as one of its products. The machine which counts the number of nails and fills the boxes with nails is suspected to be malfunctioning. Three factory workers do random testing of the number of nails in a box by each selecting 25 boxes and counting the nails. The table shown are counts from the first factory worker.

~~1970~~	~~2000~~	~~1970~~	~~2010~~	~~2040~~	~~2000~~	~~2030~~	~~2010~~	~~1990~~	~~1990~~
~~2040~~	~~2100~~	~~2030~~	~~2080~~	~~2020~~	~~2010~~	~~1890~~	~~2000~~	~~1950~~	~~2000~~
~~2010~~	~~1980~~	~~2000~~	~~2000~~	~~1960~~					

$\bar{X} = 2003.2$

$\sigma x = 40.27$

Nails	Freq
1970	II
1890	I
1950	I
1990	II
1980	I
1960	I
2000	~~IIII~~ I
2010	IIII
2040	II
2020	I
2030	II
2100	I
2080	I

Determine the mean and standard deviation, to the nearest whole number, of the counts from the first factory worker.

use 1-vars stats after puting all 25 numbers into L_1
or use 1-varstats L_1, L_2 after making a frequency table (left)

Does his data indicate that there may be a problem with the machine? Explain

Not with the average, perhaps the variation could be reduced to 20 or 50.

The second factory worker calculated a $\mu = 2100$ and $\theta = 50$. What do his results indicate?

There are too many nails on average put into each box and it varies from 2000 – 2100 95% of time.

The second factory worker calculated a $\mu = 1900$ and $\theta = 20$. What do her results indicate?

There are too few nails on average put into each box with it varying from 1880 – 1920 68% of the time or 1860 – 1940 95% of the time.

Reviewing Binomial Distributions and Confidence Intervals

1. It is estimated that 10% of the vehicles crossing the border between Canada and the U.S. contain undeclared goods. If the customs officers search 400 vehicles at random, what is the probability that:
   a. more than 50 vehicles contain undeclared goods

$\mu = np$ $\quad\quad \sigma = \sqrt{np(1-p)}$

$\mu = 400(0.10)$ $\quad\quad \sigma = \sqrt{400(0.10)(0.90)}$

$\mu = 40$ $\quad\quad \sigma = 6$

Normalcdf (50, 400, 40, 6)
$= \boxed{0.0478}$

   b. fewer than 25 vehicles contain undeclared goods.

Normalcdf (0, 25, 40, 6) $= \boxed{0.0062}$

2. The captain of a ferry estimates that 40% of the passengers will get seasick in rough weather. If, on a trip with 70 passengers, the ferry sails into rough weather, what is the probability that:

   a. more than half of the passengers will get seasick

Normalcdf (35, 70, 28, 4.1) $= \boxed{0.0439}$

$\mu = np = (0.40)(70) = 28$ $\quad\quad \sigma = \sqrt{(70)(0.4)(0.6)} = \boxed{4.1}$

b. fewer than 25 of the passengers get seasick

Normal cdf$(0, 25, 28, 4.1) = 0.2322$

3. A balloon manufacturer acknowledges that 3 % of the balloons made are defective. In a shipment of 4000 balloons, what is the probability that fewer than 100 are defective?

Normal cdf $(0, 100, 90, 9.34) = 0.8578$

$\mu = n \times p$
$= 3000 \times 0.03$
$= 90$

$\sigma = \sqrt{3000(0.03)(0.97)}$
$\sigma = 9.34$

There is an 86% probability that fewer than 100 balloons are defective.

4. A successful salesperson estimates that she can sell a car to 15 % of the potential customers who enter her dealership.

a. Estimate the number of sales that she would make in a month in which 250 potential customers enter the dealership.

$p = 0.15 \times 250 = 37.5$

$np = (250)(0.15) = 37.5$

b. Construct a 95% confidence interval for her monthly sales.

$\mu = 37.5$ $\quad n = 250$
$p = 0.15$
$(1-p) = 0.85$

$\sigma = \sqrt{250(.15)(0.85)}$
$= 5.65$

Lower
$37.5 - 1.96(5.65)$
26

Upper
$37.5 + 1.96(5.65)$
49

In 1998, Statistics Canada reported that 46 % of Canadian households use the Internet. George polled 100 households in his neighborhood to see it they use the Internet more often than the national average.

a. Construct a 95 % confidence interval for the number of households that use the Internet.

$\mu = np$
$\mu = 100(0.46)$
$\mu = 46$

$\sigma = \sqrt{np(1-p)}$
$\sigma = \sqrt{100(0.46)(0.56)}$
$\sigma = 5.07$

Lower	Upper
$46 - 1.96(5.07)$	$46 + 1.96(5.07)$
36.06	55.94

With 95% confidence George expects between 36 and 56 will use internet.

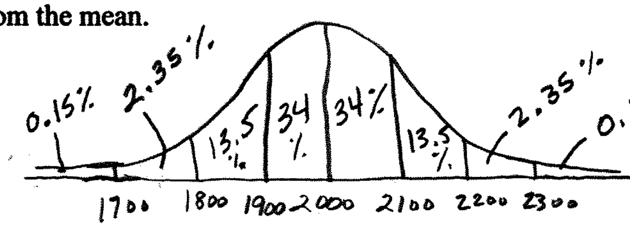
b. George found that 60 households use the Internet. Does this fall in the 95 % confidence interval?

No, it is beyond the 56 (max) expected.

Review of The Normal Distribution and z-scores

1. The mean and standard deviation of a set of data are $\mu = 2000$ and $\sigma = 100$.

a) Sketch a normal curve to show the distribution of the data. Mark the points that are 1, 2, and 3 standard deviations from the mean.



b) Between which 2 of these points does 95% of the data lie?

1800 and 2200

c) Between which 2 of these points, to the right of the mean, does 2.35% of the data lie?

2200 and 2300

d) Between which 2 of these points, to the right of the mean, does 15.85% of the data lie?

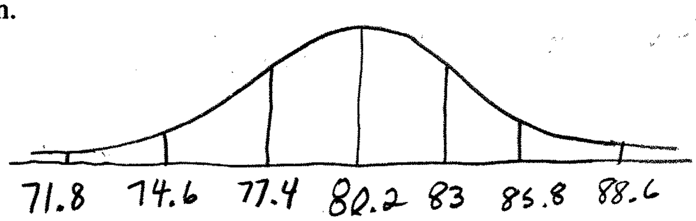
2100 and 2300

e) Two and one-half percent of the data lies to the right of which one of these points?

2200

2. The mean and standard deviation of a set of data are $\mu = 80.2$ and $\sigma = 2.8$.

a) Sketch a normal curve to show the distribution of the data. Mark the points that are 1, 2, and 3 standard deviations from the mean.



b) Between which 2 of these points does 99.7% of the data lie?

71.8 and 88.6

c) Between which 2 of these points, to the left of the mean, does 2.35% of the data lie?

71.8 and 74.6

d) Between which 2 of these points, to the left of the mean, does 15.85% of the data lie?

71.8 and 77.4

e) Sixteen percent of the data lies to the left of which one of these points?

77.4

3. Calculate the z-score for each of the following.

a) $\mu = 80$, $\sigma = 5$, and $x = 100$

$$z = \frac{100 - 80}{5} = 4$$

b) $\mu = 45.5$, $\sigma = 2.5$, and $x = 50$

$$z = \frac{50 - 45.5}{2.5} = 1.8$$

c) $\mu = 6.5$, $\sigma = 1.8$, and $x = 3$

$$z = \frac{3 - 6.5}{1.8} = -1.94$$

d) $\mu = 43.8$, $\sigma = 0.6$, and $x = 43$

$$z = \frac{43 - 43.8}{0.6} \quad z = -1.33$$