

QUANTITATIVE METHODS

Martin Huard

Winter 2005

Excel Review for Final Exam

Answer each question on a separate spreadsheet, labeling the sheet with “Question #_”. Make sure that your results are properly labeled, with relevant units.

Question 1-14 are from after the midterm, and questions 15-20 are the same questions you had on the Excel review for midterm.

1. After the private screening of a new television pilot, audience members were asked to rate the new show on a scale of 1 to 10 (10 being the highest rating). From a group of 140 people, the accompanying responses were obtained.

Rating	1	2	3	4	5	6	7	8	9	10
Frequency	1	4	3	11	23	21	28	29	16	4

- a) Construct a probability distribution table for the rating of a pilot.
 - b) Draw a histogram.
 - c) What is the mean score?
 - d) What is the standard deviation?
2. Consider the normal distribution with $m=100$ and $s=16$.
 - a) Find $P(x < 110)$.
 - b) Find $P(x > 115)$
 - c) Find $P(80 < x < 120)$.
 - d) Find the value of x such that the area under the curve to the left of x is 0.80.
 - e) Find the value of x such that the area under the curve to the right of x is 0.10.
 3. Express Courier Service has found that the delivery times for packages are normally distributed with mean 14 hours and standard deviation 2 hours.
 - a) For a package selected at random, what is the probability that it will be delivered in 18 hours or less?
 - b) For a package selected at random, what is the probability that it will be delivered in more than 15 hours?
 - c) For a package selected at random, what is the probability that it will be delivered in between 10 and 20 hours?
 - d) What should the guaranteed delivery time on all packages be in order to be 95% sure that a given package will be delivered within this time?

4. Do people who stop smoking tend to gain weight? A study (later published in a medical journal) was undertaken to answer this question. The authors of this study collected data on a random sample of 315 men over the age of 35 who had quit smoking during the past 10 years and found that these men had gained an average of 5.28 kilograms since quitting smoking with a standard deviation of 0.59 kilogram. Construct a 98% confidence interval for the corresponding population mean.
5. How large are math classes in CEGEP's? A random sample of 100 math classes was taken and the number of students in each class counted. Construct a 92% confidence interval for the mean class size for all CEGEP math classes. The data can be found on the worksheet "Data – Math Classes" from my web site.
6. To determine the views of students at SLC on whether an extremist hate group should be given a permit to demonstrate, a seven-point attitude scale (1 = strongly opposed through 7 = strongly favor) was administered to a random sample of 15 students. This survey yielded a sample mean of 2.1 and a standard deviation of 1.5. Assuming that the scores are approximately normally distributed, construct a 95% confidence interval for the mean population score.
7. Suppose a researcher wanted to examine the extent of cooperation in kindergarten children. To do so, she unobtrusively observes a group of children at play for 30 minutes and notes the number of cooperative acts engaged in by each child. Here are the number of cooperative acts exhibited by each child:

1 5 2 3 4 1 2 2 4 3

Construct a 90% confidence interval for the mean number of cooperative acts exhibited by children, assuming that the number of cooperative acts exhibited by children is normally distributed.

8. A random sample of 40 workers who log on to Yahoo while at work were asked how long they stayed logged on while at work. The mean length was 106 minutes with a standard deviation of 42 minutes. Using this sample, can you conclude that the mean time spent on Yahoo is greater than 90 minutes at the 3% level of significance?
9. The minister of education claims that college math classes have fewer than 30 students. A random sample of 100 college math classes was taken and the number of students in each class counted. At the 5% level of significance, can you conclude that the minister's claim is valid? The data can be found on the worksheet "Data – Math Classes" from my web site.
10. A sample of 225 parents were asked how much time they spent per week on school work or school-related activities. This sample produced a mean of 5.6 hours per week, with a standard deviation of 4.4 hours. At the 1% level of significance, test the claim that the mean number of hours spent by parents on school work or school-related activities is 5 hours per week.

11. At the 1% level of significance, test the claim that the hours worked by college students is greater than 15 hours per week. A random sample of 25 students produced a sample mean of 20.83 hours per week with a standard deviation of 14.20 hours per week. Assume that the number of hours worked by college students is normally distributed.
12. According to the last census on the Canadian population, the mean age of Quebecers is 38.8 years. A random sample of 15 residents from Ste-Foy revealed that they had a mean age of 42.2 years with a standard deviation of 11.9 years. Is this sufficient evidence, at the 5% level of significance, to conclude that the average age of downtown Quebec city residents is different then that of the general Canadian population? Assume that ages are normally distributed.
13. During a television miniseries, what is the average length of time between commercial breaks? A random sample of 15 such periods was selected from miniseries that were aired on commercial television stations last year. The times between commercial breaks were can be found on the worksheet “Data – Time Between Breaks” from my web site. At the 3% level of significance, can you conclude that the average time between commercial breaks is less than 12 minutes? Assume that the length of time between commercial breaks is normally distributed.
14. A random sample of 1000 adults was asked whether they voted in the last general election. Here are the results, broken down by age group.

	18-24 years	25-39 years	40-59 years	60 years or older
Voted	46	159	178	159
Did not vote	90	179	118	71

At the 1% level of significance, test the claim that voting status and are age independent.

15. In a survey on Canadian education, a random sample of 500 students rated their teacher. The results were as follows: 123 rated their teacher as Excellent, 225 as Very Good, 133 as Good and 19 as Not Good.
- With these results, construct a relative frequency distribution table.
 - Sketch a Pie Chart
16. A snack size bag of peanut M&M candies contains 21 candies of different colors.
- Open the results of the color count in the file “Data - Mid-Term Review”.
 - Construct a relative frequency table, letting Excel do the counting.
 - Sketch a Pareto diagram
17. Sketch a time plot on the monthly amount of sunshine in Quebec City. The data can be found in the file “Data - Mid-Term Review”.

18. The test scores on a 100-point test were recorded for a class of 20 students.

61	93	91	55	63	86	82	76	57	76
94	89	67	62	72	87	68	65	75	84

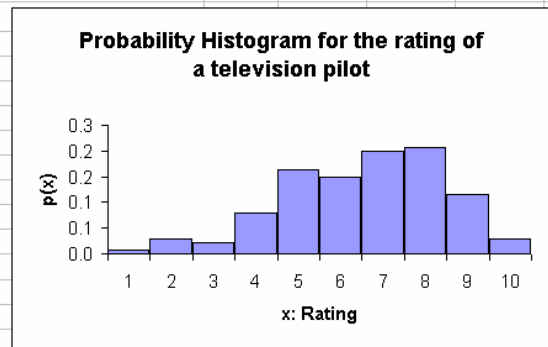
- Enter the data in a column.
 - Find the mean, median and mode, range, variance and the standard deviation.
 - Find the 1st and 3rd quartiles and the 90th percentile.
19. For her geography project, Sara collected the average wind speeds (in kilometer per hour) for 45 selected cities in Canada.
- Open the results for her project in the file “Data-Mid-Term Review” and copy it on your spreadsheet.
 - Compile the results in the form of a relative frequency distribution table, letting Excel do the counting.
 - Sketch a histogram from the table obtained in (b)
20. If you try to rent an apartment or buy a house, you find that many real estate representatives establish apartment rents and house prices on the basis of square footage of heated floor space. The data in the table give the square footages and sales prices for a number of randomly selected houses from those sold in a small city.
- Open the table in the file “Data - Mid-Term Review” and copy it on your spreadsheet.
 - Find the equation of the least-squares line.
 - Find the coefficient of correlation and the coefficient of determination.
 - If a house has an area of 1550 square feet, what is the forecasted price?
 - Sketch a scatter diagram containing the least-squares line.

Answers

Note: The shaded numbers were found using formulas.

Question 1

3	Rating	Frequency	x	p(x)	xp(x)	x ² p(x)
4	1	1	1	0.0071	0.0071	0.0071
5	2	4	2	0.0286	0.0571	0.1143
6	3	3	3	0.0214	0.0643	0.1929
7	4	11	4	0.0786	0.3143	1.2571
8	5	23	5	0.1643	0.8214	4.1071
9	6	21	6	0.1500	0.9000	5.4000
10	7	28	7	0.2000	1.4000	9.8000
11	8	29	8	0.2071	1.6571	13.2571
12	9	16	9	0.1143	1.0286	9.2571
13	10	4	10	0.0286	0.2857	2.8571
14		140		1.0000	6.5357	46.2500
15						
16						
17			Mean =	6.54		
18			Standard Deviation =	1.88		
19						



Question 2

3		mean =	100
4		standard deviation =	16
5			
6	a)	$P(x < 110) =$	0.7340
7	b)	$P(x > 115) =$	0.1743
8	c)	$P(80 < x < 120) =$	0.7887
9	d)	x =	113.5
10	e)	x =	120.5

Question 3

5		mean =	14	hours
6		standard deviation =	2	hours
7				
8	a)	$P(x < 18) =$	0.9772	
9	b)	$P(x > 15) =$	0.3085	
10	c)	$P(10 < x < 20) =$	0.9759	
11	d)	x =	17.29	hours
12				

Question 4

5	Assumptions: n =	315	≥ 30
6	Sample Mean =	5.28	kg
7	Sample St. Dev =	0.59	kg
8			
9	c =	98%	
10	E =	0.077	
11			
12	98% confidence interval:	5.20	kg to 5.36

Question 5

5	# students		
6	26	Assumptions: n =	100 ≥ 30
7	19	Sample Mean =	27.7 students
8	30	Sample St. Dev =	7.73 students
9	37		
10	37	c =	92%
11	40	E =	1.353
12	12		
13	26	92% confidence interval:	26.34 students to 29.04 students
14	30		

Question 6

4			
5	Assumptions: Population is normally distributed		
6	n =	15	
7	Sample Mean =	2.1	
8	Sample St. Dev =	1.5	
9			
10	c =	95%	
11	t _c =	2.145	
12	E =	0.83	
13			
14			
15	95% confidence interval:	1.27	to 2.93

Question 7

4									
5	# cooperative acts	Method 1							
6	1	Assumptions:	Population is normally distributed						
7	5	n =	10						
8	2	Sample Mean =	2.7 cooperative acts						
9	3	Sample St. Dev =	1.34 cooperative acts						
10	4								
11	1	c =	90%						
12	2	t _c =	1.833						
13	2	E =	0.775						
14	4								
15	3	90% confidence interval:	1.92 cooperative acts	to	3.48 cooperative acts				
16									
17		Method 2							
18		# cooperative acts							
19									
20		Mean	2.7						
21		Standard Error	0.4229526						
22		Median	2.5						
23		Mode	2						
24		Standard Deviation	1.3374935						
25		Sample Variance	1.7888889						
26		Kurtosis	-0.851654						
27		Skewness	0.3343605						
28		Range	4						
29		Minimum	1						
30		Maximum	5						
31		Sum	27						
32		Count	10						
33		Confidence Level(90.0%)	0.7753202						
34									
35		92% confidence interval:	1.92 cooperative acts	to	3.48 cooperative acts				

Question 8

4	Assumptions: n =	40	≥ 30						
5	Ho : μ =	90 min							
6	Ha : μ >	90 min							
7									
8	Right-tailed Test								
9	α =	3%							
10	z critical =	1.88							
11									
12	Sample Mean =	106 min							
13	Sample St. Dev =	42 min							
14									
15	z =	2.41		Or, using the p-value:	p-value =	0.0080			
16				z is in the critical region					
17	Decision:	Reject Ho							
18									
19	Therefore there is sufficient evidence at the 3% level of								
20	significance to conclude that that the mean time workers spent								
21	on Yahoo is greater than 90 minutes.								

Question 9

5	# students								
6	26	Assumptions: n =	100	≥ 30					
7	19	Ho : μ =	30 students						
8	30	Ha : μ <	30 students						
9	37								
10	37	Left-tailed Test							
11	40	α =	5%						
12	12	z critical =	-1.645						
13	26								
14	36	Sample Mean =	27.69 students						
15	20	Sample St. Dev =	7.73 students						
16	23								
17	16	z =	-2.988		Or, using the p-value:	p-value =	0.0014		
18	15			z is in the critical region					
19	21	Decision:	Reject H0						
20	23								
21	13	Therefore there is sufficient evidence at the 5%							
22	24	level of significance to conclude that college math							
23	25	classes have fewer than 30 students.							
24	29								

Question 10

4	Assumptions: n =	225	≥ 30		
5	Ho : μ =	5	hours		
6	Ha : μ ≠	5	hours		
7					
8	Two-tailed Test				
9	α =	1%			
10	z critical =	2.58			
11					
12	Sample Mean =	5.6	hours		
13	Sample St. Dev =	4.4	hours		
14				Or, using the p-value:	
15	z =	2.05		p-value =	0.0408
16				z is not in the critical region	
17	Decision:	Fail to reject Ho			
18					
19	Therefore there is not sufficient evidence at the 1% level of				
20	significance to conclude that the mean number of hours spent by				
21	parents on school work or school-related activities is not 5 hours				
22	per week.				
23					

Question 11

5	Assumptions: Population is normally distributed				
6	n =	25			
7	Ho : μ =	15	hours		
8	Ha : μ >	15	hours		
9					
10	Right-tailed Test				
11	α =	1%			
12	t critical =	2.492			
13					
14	Sample Mean =	20.83	hours		
15	Sample St. Dev =	14.2	hours		
16					
17	t =	2.05			
18				t is not in the critical region	
19	Decision:	Fail to reject H0			
20					
21	Therefore there is not sufficient evidence at the 5% level of significance				
22	to conclude that the number of hours worked by college students is				
23	greater than 15 hours per week.				

Question 12

5	Assumptions: Population is normally distributed				
6	n =	15			
7	Ho : μ =	38.8	hours		
8	Ha : μ ≠	38.8	hours		
9					
10	Two-tailed Test				
11	α =	5%			
12	t critical =	2.145			
13					
14	Sample Mean =	42.2	hours		
15	Sample St. Dev =	11.9	hours		
16					
17	t =	1.11			
18				t is not in the critical region	
19	Decision:	Fail to reject H0			
20					
21	Therefore there is not sufficient evidence, at the 5% level of significance,				
22	to conclude that the average age of downtown Quebec city residents is				
23	different then that of the general Canadian population.				

Question 13

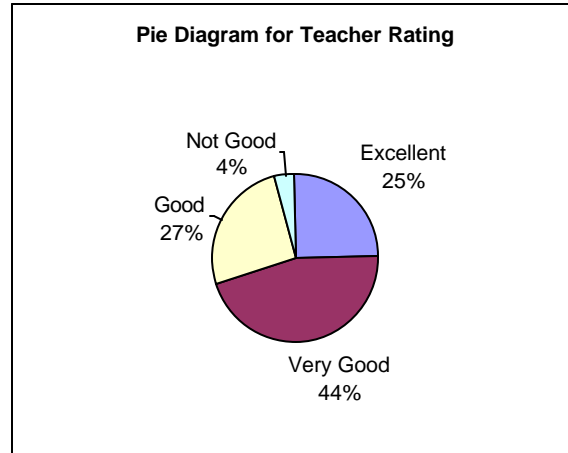
4	Time (min)				
5	5	Assumptions: Population is normally distributed			
6	12	n =	16		
7	8	Ho : μ =	12	min	
8	14	Ha : μ <	12	min	
9	13				
10	10	Left-tailed Test			
11	9	α =	3%		
12	8	t critical =	-2.034		
13	14				
14	11	Sample Mean =	10.06	min	
15	9	Sample St. Dev =	2.62	min	
16	10				
17	6	t =	-2.96		
18	9			t is in in the critical region	
19	12	Decision:	Reject H0		
20	11				
21		Therefore there is sufficient evidence, at the 5% level			
22		of significance, to conclude that the average time			
23		between commercial breaks is less than 12 minutes.			
24					

Question 14

5	Observed Values										
6		18-24 years	25-39 years	40-59 years	60 years or older		Assumptions: Classes are mutually exclusive and all inclusive				
7	Voted	46	159	178	159	542					
8	Did not vote	90	179	118	71	458					
9		136	338	296	230	1000					
10											
11											
12	Expected Values										
13		18-24 years	25-39 years	40-59 years	60 years or older		Decision : Reject Ho				
14	Voted	73.712	183.196	160.432	124.66	542					
15	Did not vote	62.288	154.804	135.568	105.34	458					
16		136	338	296	230	1000					
17											
18	Chi-Square Calculations										
19		10.418	3.196	1.924	9.460		p-value =	0.0000			
20		12.329	3.782	2.277	11.195						
21											

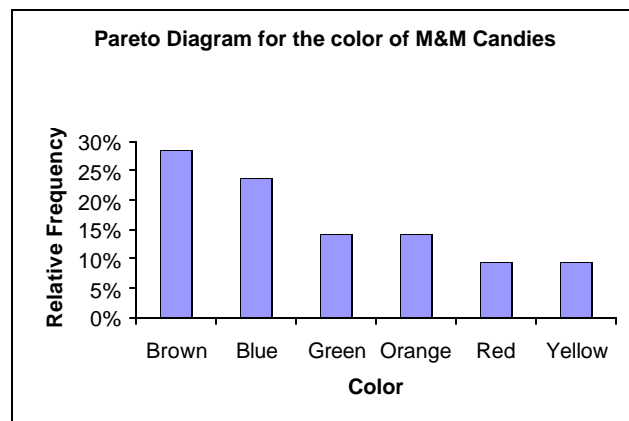
Question 15

Rating	Frequency	Relative Frequency
Excellent	123	25%
Very Good	225	45%
Good	133	27%
Not Good	19	4%
	500	100%

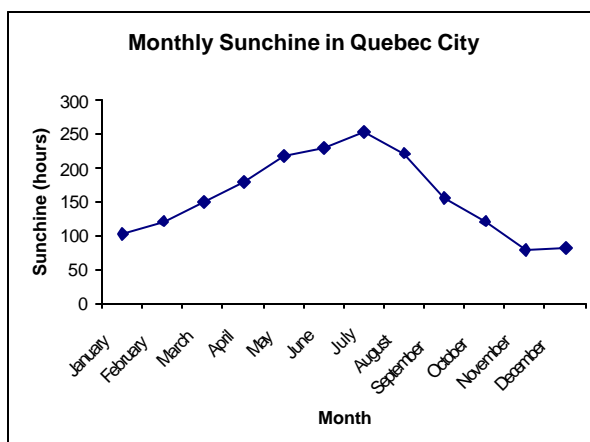


Question 16

Color	Frequency	Relative Frequency
Brown	6	29%
Green	3	14%
Blue	5	24%
Red	2	10%
Yellow	2	10%
Orange	3	14%
	21	100%



Question 17



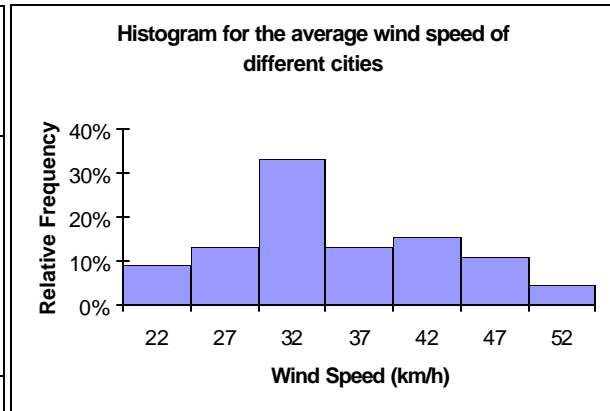
Question 18

Measures for the test-scores in a class

Mean	75.15
Median	75.5
Mode	76
Range	39
Variance	149.4
Standard Deviation	12.22
Q ₁	64.5
Q ₂	86.25
P ₉₀	91.2

Question 19

Wind Speed (km/h)	Frequency	Relative Frequency	Midpoint
20 to 24	4	9%	22
25 to 29	6	13%	27
30 to 34	15	33%	32
35 to 39	6	13%	37
40 to 44	7	16%	42
45 to 49	5	11%	47
50 to 54	2	4%	52
<i>Total</i>	45	100%	



Question 20

Regression and Correlation for Area vs Price of housing

Equation of least-squares line $y = 51\,206 + 27.41x$

Coefficient of correlation $r = 0.978$

Coefficient of determination $r^2 = 0.9574$

Forecast price for an area of \$ 93 685

1550 square feet

