Solutions for Practice Questions for Exam 3 **Math 283**

1. Assume that sample data, based on two independent samples of size 25, gives us $\bar{x}_1 = 505$, $\bar{x}_2 = 515$, $s_1 = 23$, and $s_2 = 28$.
   a. What is a 95% confidence interval (use the conservative value for the degrees of freedom) for $\mu_2 - \mu_1$?

      (–4.57, 24.57)

   b. Determine whether each of the following statements is true or false.
      i. True_ Based on the confidence interval, we can conclude, at the 5% significance level, that there is no difference between the two population means, μ_2 and μ_1.
      ii. True_ The margin of error for the difference between the two sample means would be smaller if we were to take larger samples.
      iii. False_ If we were to use the unpooled *t* test with the more accurate approximation for the degrees of freedom (used in software), the degrees of freedom would be 51.
      iv. True_ If a 99% confidence interval were calculated instead of the 95% interval, it would include more values for the difference between the two population means.

2. Some researchers have conjectured that stem-pitting disease in peach tree seedlings might be controlled with weed and soil treatment. An experiment is conducted to compare peach tree seedling growth when the soil and weeds are treated with one of two herbicides. In a field containing 20 seedlings, 10 are randomly selected throughout the field and assigned to receive Herbicide A. The remainder of the seedlings is assigned to receive Herbicide B. Soil and weeds for each seedling are treated with the appropriate herbicide, and at the end of the study period the height in centimeters is recorded for each seedling. The following results are obtained:

Herbicide A	Herbicide B
$\bar{x}_1 = 94.5$ cm	$\bar{x}_2 = 109.1$ cm
$s_1 = 10$ cm	$s_2 = 9$ cm

   a. What is a 90% confidence interval (use the conservative value for the degrees of freedom) for $\mu_2 - \mu_1$?
      i. 14.6 ± 7.00
      ii. 14.6 ± 7.38
      iii. 14.6 ± 7.80
      iv. 14.6 ± 9.62

   b. Suppose we wish to determine if there tends to be a difference in height for the seedlings treated with the different herbicides. To answer this question, we decide to test the hypotheses H_0: $\mu_2 - \mu_1 = 0$, H_a: $\mu_2 - \mu_1 \neq 0$. What is the value of the two-sample *t* statistic?
      i. 14.60
      ii. 7.80
      iii. 3.43
      iv. 2.54

c. What can we say about the value of the *P*-value?
   i. **P-value < 0.01**
   ii. $0.01 < P\text{-value} < 0.05$
   iii. $0.05 < P\text{-value} < 0.10$
   iv. P-value > 0.10

3. A simple random sample of 120 vet clinics in the Midwest reveals that the vast majority of them only treat small pets (dogs, cats, rabbits, etc.) and no large animals (cows, horses, etc.). Of the 120 clinics sampled, 88 responded that they do not treat large animals at their clinic.

   a. What is the value of the standard error of $\hat{p}$?
      i. 0.02
      ii. 0.03
      iii. 0.04
      iv. 0.05

   b. What is a 90% confidence interval for p, the population proportion of vet clinics that *do* treat large animals?
      i. (0.163, 0.371)
      ii. (0.188, 0.346)
      iii. (0.200, 0.333)
      iv. (0.667, 0.800)

   c. If a 95% confidence interval were calculated instead, what would happen to the width of the confidence interval?
      i. It would be narrower.
      ii. It would stay the same.
      iii. It would be wider.
      iv. This cannot be determined from the information given.

4. At a large Midwestern university, a simple random sample of 100 entering freshmen in 1993 found that 20 of the sampled freshmen finished in the bottom third of their high school class. Admission standards at the university were tightened in 1995. In 1997, a simple random sample of 100 entering freshmen found that only 10 finished in the bottom third of their high school class. Let p_1 and p_2 be the proportions of all entering freshmen in 1993 and 1997, respectively, who graduated in the bottom third of their high school class.

   a. Is there evidence that the proportion of freshmen who graduated in the bottom third of their high school classes in 1997 has been reduced, as a result of the tougher admission standards adopted in 1995, compared to the proportion in 1993? To determine this, test the hypotheses H_0: $p_1 = p_2$ versus H_a: $p_1 > p_2$. What is the value of the z statistic for testing these hypotheses?
      i. $z = 1.20$
      ii. $z = 1.92$
      iii. $z = 1.96$
      iv. $z = 1.98$

b. What is the value of the *P*-value?
   i. *P*-value < 0.001
   ii. *P*-value = 0.0238
   iii. *P*-value = 0.025
   iv. *P*-value = 0.1151

5. A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the table below.

Managers			Workers		
	Gender			Gender	
Income	Male	Female	Income	Male	Female
$100,000	80	20	$10,000	30	20
$200,000	20	30	$20,000	20	80

a. What is the proportion of male managers who make $200,000 per year?
   i. 0.067
   ii. 0.133
   iii. 0.2
   iv. 0.4

b. What is the proportion of female managers who make $200,000 per year?
   i. 0.1
   ii. 0.2
   iii. 0.4
   iv. 0.6

c. What proportion of the managers is female?
   i. 0.2
   ii. 0.333
   iii. 0.5
   iv. 0.667

d. What conclusion(s) can we draw from this table?
   i. The mean salary of female managers is greater than that of male managers.
   ii. The mean salary of males in this business is greater than the mean salary of females.
   iii. The mean salary of female workers is greater than that of male workers.
   iv. All of the above.

6. Could mudwrestling be the cause of a rash contracted by University of Washington students in the spring of 1992? Two physicians at the University of Washington student health center wondered this when one male and six female students complained of rashes after participating in a mud-wrestling event. Questionnaires were sent to all students in the residence halls who participated in the event. The questionnaire asked about the appearance of a rash and about attendance at the mud-wrestling event. The results, by gender, are summarized in the following table.

	Gender	
Developed rash?	Female	Male
Yes	12	12
No	12	38

The SPSS output for the above table is given below. The output includes the cell counts, the expected cell counts, the chi-square statistic, and its degrees of freedom. Expected counts are printed below observed counts.

Developed Rash? * Gender Crosstabulation

			Gender		Total
			Female	Male	
Developed Rash?	Yes	Count	12	12	24
		Expected Count	7.78	16.22	24.0
	No	Count	12	38	50
		Expected Count	16.22	33.78	50.0
Total		Count	24	50	74
		Expected Count	24.0	50.0	74.0

Chi-square = 2.289 + 1.098 + 1.098 + 0.527 = 5.002 df = 1, *P*-value = 0.0253

a. Which cell contributes most to the chi-square statistic?
   i. Men who developed a rash.
   ii. Men who did not develop a rash.
   iii. Women who developed a rash.
   iv. Women who did not develop a rash.

b. What conclusion can we draw from the above tables at the 5% significance level?
   i. There appears to be evidence of an association between the gender of an individual attending the event and development of a rash.
   ii. Mud wrestling causes a rash, especially for women.
   iii. There is absolutely no evidence of any relationship between the gender of an individual attending the event and the subsequent development of a rash.
   iv. Development of a rash is a real possibility if you attend mud-wrestling events, especially if you attend them on a regular basis.

7. A simple random sample of 100 college students was interviewed. They were asked what size pizza they usually order and what there favorite topping is. The results are presented below.

	Topping			
Size	Pepperoni	Veggie	Cheese	Total
Small	18	11	6	35
Medium	14	12	7	33
Large	3	9	20	32
Total	35	32	33	100

a. What would be null hypothesis for a chi-square test based on these data?
   i. Pizza topping and pizza size are independent.

ii. The average pizza size is the same for pepperoni, veggie, and cheese pizzas.
iii. The distribution of pizza topping is the same for small, medium, and large pizzas.
iv. The distribution of pizza size is different for the three different pizza toppings.

b. Under the null hypothesis that there is no association between pizza size and pizza topping, what is the value of the expected count for a small pepperoni pizza?
i. 11.67
ii. 12
iii. 12.25
iv. 13

c. What are the appropriate degrees of freedom for the chi-square statistic?
i. 3
ii. 4
iii. 6
iv. 8

d. What is the contribution to the chi-square statistic from the cell of a large cheese pizza?
i. 8.4
ii. 9.3
iii. 20.4
iv. 52.8

e. Are the data statistically significant at the 10% significance level? Hint: You do not need to calculate the actual value of the chi-square statistic to answer this question.
i. Yes
ii. No
iii. This cannot be determined from the information given.

8. Deborah's Dairy Market sells both cottage cheese and ice cream. The sales manager recently noticed that in months when the company sells more cottage cheese, it seems to sell more ice cream as well, on average. One of his aides was assigned to put this theory to the test. The aide's analysis of sales data for the past twelve months (in millions of pounds for both the cottage cheese and the ice cream) gives the following SPSS output.

Descriptive Statistics

	Mean	Std. Deviation	N
ICECREAM	63.90	9.53	12
COTTAGE	76.63	5.10	12

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.611	.373	.310	7.92

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-23.67	35.98		-.66	.526
	COTTAGE	1.14	.47	.611	2.44	.035

a. What is the equation of the least squares regression line?
$\hat{y} = -23.67 + 1.14\, x$

b. Suppose the aide wishes to test the hypotheses H_0: $\beta_1 = 0$ versus H_a: $\beta_1 > 0$. What is the value of the *P*-value for this test?
P-value = 0.0175

c. Fill in the blank. At the 5% significance level, the aide should report to the sales manager that the mean of ice cream sales (select one) **does** / does not appear to increase as the cottage cheese sales increase.

d. What is an (approximate) 95% confidence interval for the slope β_1?
(0.093, 2.19)

e. Cottage cheese sales for the next month are 82 million pounds. Ice cream sales are not yet available. Predict how much ice cream Deborah's Dairy Market will sell in this month with a 95% interval. Express the interval in the appropriate units.
(50.61 million lbs., 89.01 million lbs.)

f. The following (partial) ANOVA table was obtained from statistical software.

Source	DF	Sum of Squares
Model	1	372.97
Error		627.08

i. What are the degrees of freedom for the MSE?
10

ii. What is the value of the MSE?
62.7

iii. What is the value of SST, the total sum of squares?
1000.05

iv. What is the value of the *F*-statistic for testing the hypotheses H_0: $\beta_1 = 0$ versus H_a: $\beta_1 \neq 0$?
$F = 5.95$