Math 205
Spring 2015
Dr. Lily Yen

Midterm 2

Show all your work

Name: ____________________

Score: ___/40

Problem 1: Lily's Restaurant serves three fixed-price dinners costing \$12, \$15, and \$20. For a randomly selected couple dining at Lily's, let X be the cost of the woman's dinner and Y be the cost of the man's dinner. The joint pmf of X and Y is given in the following table:

$p(x,y)$		y			
		12	15	20	$p_X(x)$
	12	0.05	0.05	0.00	0.10
x	15	0.05	0.10	0.20	0.35
	20	0.10	0.35	0.10	0.55
$p_Y(y)$		0.20	0.50	0.30	

a. Compute the marginal pmf's of X and Y, i.e., $p_X(x)$ and $p_Y(y)$. You may use the space around the table to help you.

See the table above.

Score: /2

b. Find $P(X \leq 15 \text{ and } Y \leq 15)$. Describe the probability in words.

The probability that both parties select a meal that costs at most \$15 is $0.05 + 0.05 + 0.05 + 0.10 = 0.25$

Score: /2

c. Are X and Y independent? Justify your answer.

The two random variables are *not independent* since, for example, $p_X(12) \times p_Y(12) = 0.10 \times 0.20 = 0.02 \neq 0.05 = p(12, 12)$. You can also pick other points like $(15, 15)$ to prove dependence.

Score: /2

d. What is the expected total cost of dinner for a couple? Show how you obtain your answer.

$(12+12)0.05 + (12+15)0.05 + (12+20)0.00 + (15+12)0.05 + (15+15)0.10 + (15+20)0.20 + (20+12)0.10 + (20+15)0.35 + (20+20)0.10 = \$33.35$.

Score: /2

Problem 2: Hastings, Cassiar, and Mcgill streets feed into Iron Workers' Memorial Bridge entrance. Suppose that between 7 am and 7:30 am on a weekday morning, the number of vehicles coming from each road onto the bridge is a random variable, with expected value and standard deviation as given in the table.

	Hastings	Cassiar	Mcgill
Expected value:	800	1000	600
Standard deviation:	16	25	18

a. What is the expected total number of vehicles entering the bridge at this point during the period?

$E(\sum X) = 800 + 1000 + 600 = 2400$.

Score: /1

b. What is the variance of the total number of entering vehicles? Specify any assumption made about the relationship between the number of vehicles on the different streets.

If the random variables are *independent*, then the variance of the sum is the sum of the variances: $V(\sum X) = 16^2 + 25^2 + 18^2 = 1205$, so the standard deviation is $\sqrt{1205} \approx 34.71$.

Score: /2

Problem 3: Capilano U Publishing House is interested in estimating the strength of the bindings produced by a particular binding machine. Strength can be measured by recording the force required to pull the pages from the binding. If this force is measured in pounds, how many books should be tested to estimate the average force required to break the binding to within 0.1 pound with 95% confidence? Assume that σ is known to be 0.8.

A 95% confidence interval has endpoints $\bar{x} \pm 1.96 \cdot \frac{\sigma}{\sqrt{n}}$. Here $\sigma = 0.8$, and $1.96 \cdot \frac{\sigma}{\sqrt{n}} = 0.1$, so $n = (\frac{1.96 \cdot \sigma}{0.1})^2 = 246$ books.

Score: /3

Problem 4: An Ironman competition consisting of swimming, cycling, and running is one of the more strenuous amateur sporting events. A research article reports on a study involving 9 male participants. Maximum heart rate (beats/min) was recorded during the performance of each of the three events. For swimming, the sample mean and sample standard deviation were 188.0 beats/min and 7.2 beats/min, respectively. Assuming that the heart-rate distribution is (approximately) normal, construct a 98% CI for true mean heart rate of Ironman participants while swimming.

- Name the test required for the confidence interval.
- State all parameters with their corresponding values.
- All calculator usage must be stated with its results listed.
- Summarize giving 2-decimal place accuracy with a complete sentence regarding reliability and precision.

Here sample size is $n = 9$, sample mean $\bar{x} = 188.0$, and sample standard deviation $s = 7.2$. We need t test because of a small sample with $\alpha = 0.02$ and the degree of freedom, $\nu = 8$. In TI-83 or 84, use inverse T program to find the value of $t_{0.01,8} = 2.896459$ to complete the calculations below.
The 98% confidence interval (so $\alpha = 0.02$) for μ is
$\left(\bar{x} - t_{\alpha/2,n-1} \cdot \frac{1}{\sqrt{n}} s, \bar{x} + t_{\alpha/2,n-1} \cdot \frac{1}{\sqrt{n}} s\right) = \left(188.0 - t_{0.01,8} \cdot \frac{1}{3} \cdot 7.2, 188.0 + t_{0.01,8} \cdot \frac{1}{3} \cdot 7.2\right) =$
$\left(188.0 - 2.8965 \cdot \frac{1}{3} \cdot 7.2, 188.0 + 2.8965 \cdot \frac{1}{3} \cdot 7.2\right) = (181.05, 194.95)$.
The true mean heart rate in beats per minute, μ, is contained in the interval $(181.05, 194.95)$ 49 times out of 50.

Score: /5

Problem 5: Minor surgery on horses under field conditions requires a reliable short-term anesthetic producing good muscle relaxation, minimal cardiovascular and respiratory changes, and a quick, smooth recovery with minimal aftereffects so that horses can be left unattended. A research article reports that for a sample of 73 horses to which ketamine was administered under certain conditions, the sample average lying-down time was 18.86 min and the standard deviation was 8.6 min.

a. Does this data suggest that true average lying-down time under these conditions is less than 20 minutes?

- Use a significance level of 0.10.
- Define your variables and identify parameters.
- State hypotheses.
- Name the test and your test statistic.
- Shade the corresponding rejection region to your test with its value.
- State your conclusion with a complete sentence.

Let μ be true population mean in minutes; $n = 73$, sample size; $\bar{x} = 18.86$ minutes, sample mean; and $\sigma = 8.6$ minutes, standard deviation.
Here the null hypothesis is $H_0 : \mu = 20$ and the alternative hypothesis is $H_\mathrm{a} : \mu < 20$.
Since $n = 73 > 40$, you can use z-test and that the test statistic: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ approximately follows a standard normal distribution if H_0 is true. Now $z = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{18.86-20}{8.6/\sqrt{73}} = -1.133$, and normalcdf$(-1 \times 10^{99}, -1.133) = 0.1287 > 0.10$, so you cannot reject H_0.
(So the available data does not prove that the average lying-down time is less than 20 min.)

b. Define a Type I error.

A Type I error consists of rejecting the null hypothesis H_0 when it is true.

c. Find Type II error to 3 decimal places when the true mean is 21 minutes.

A Type II error involves not rejecting H_0 when H_0 is false. For this particular problem,

$$\beta(21) = 1 - \Phi(-z_{0.1} + \frac{20-21}{8.6/\sqrt{73}}) \approx 0.989$$

Score: /8

Problem 6: As the population ages, there is increasing concern about accident-related injuries to the elderly. A research paper reported on an experiment in which the maximum lean angle—the furthest a subject is able to lean and still recover in one step—was determined for both a sample of younger females (21–29 years) and a sample of older females (67–81 years). The following observations are consistent with summary data given in the article:

YF:	29,	34,	33,	27,	28,	32,	31,	34,	32,	27
OF:	18,	15,	23,	13,	12					

Does the data suggest that true average maximum lean angle for older females is more than 10 degrees smaller than it is for younger females? State and test the relevant hypotheses at significant level 0.10.

a. Define your variables and identify parameters.
b. State hypotheses.
c. Name the test and your test statistic.
d. Write both input and output values from your calculator.
e. State your conclusion with a complete sentence.

Let μ_1 and μ_2 be the true average maximum lean angle for young females (X_1) and old females (X_2) respectively. We are asked to test whether $\mu_1 - \mu_2 > 10$ to check if μ_2 is more than 10 degrees smaller than μ_1, or that the difference between their true averages is more than 10 degrees. As usual, we use $\bar{x}_1$ and $\bar{x}_2$ to denote sample averages.
$H_0 : \mu_1 - \mu_2 \leq 10$ vs. $H_a : \mu_1 - \mu_2 > 10$.
The samples are small, so use the two-sample t-test. Store the young females' lean angles in L_1, the older females' lean angles in L_3, and store $L_3 + 10$ in L_2. This changes the variables in the calculator where our μ_1 is still μ_1 in TI, but our $\mu_2 + 10$ is now μ_2 in TI.
Our test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 10}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

with $m = 10$ and $n = 5$.
Then $\boxed{\text{STAT}}$ TESTS 4:2-SampTTest under DATA option yields

$$\mu_1 > \mu_2, \qquad t = 2.076 \qquad p = 0.043\,27 \qquad df = 5.592$$

From the output, we round down degrees of freedom to $\nu = 5$, and apply an upper-tailed test by using inverse T program to find $t_{0.1,5} \approx 1.475884$. Since the t value from our data $t = 2.076431672$ lies in the rejection region, that is greater than 1.475884, we reject the null hypothesis.
We conclude that at the significance level of 0.10, there is strong evidence that the average maximum lean angle for older females is more than 10 degrees smaller than it is for younger females.

Score: /6

Problem 7: An article reports the following data on total iron for four types of iron formation: carbonate, silicate, magnetite, and hematite.

carbonate:	20.5	28.1	27.8	27.0	28.0
silicate:	26.3	24.0	26.2	20.2	23.7
magnetite:	29.5	34.0	27.5	29.4	27.9
hematite:	36.5	44.2	34.1	30.3	31.4

Carry out an analysis of variance F test at significance level 0.01, and summarize the results in an ANOVA table.

$\boxed{\text{STAT}}$ TESTS H:ANOVA(

Source of Variation	df	Sum of squares	Mean Square	f	p
Treatments	3	358.1	119.4	8.836	0.001 097
Error	16	216.1	13.51		
Total	19	574.2	132.9		

a. Define your variables and identify parameters.

b. State hypotheses.

c. What is the definition of the test statistic f?

d. How many treatments does the report contain?

e. How many observations does each treatment have?

f. Find the explicit value of the test statistic at $\alpha = 0.01$.

Let I be the number of treatments, J be the number of observed values in each treatment, and $MSTr$ be treatment mean square, MSS be error mean square. Let μ_1, μ_2, μ_3, μ_4 be the true mean of carbonate, silicate, magnetite, and hematite iron formations respectively corresponding to sample means $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, and $\bar{x}_4$.
Store corresponding data set in TI under L_1, L_2, L_3, and L_4.
Then $I = 4$, $J = 5$; $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ whereas $H_a : \mu_i \neq \mu_j$ for some $i \neq j$ and $i, j \in \{1, 2, 3, 4\}$.
We need to use F distribution where $F = MSTr/MSS$. This is shown in the table under f.
The report contains 4 treatments and 5 observed values for each treatment.
Graph $Y_1 = \text{Fcdf}(0, x, 3, 16)$ and $Y_2 = 1 - 0.01$ to find the intersection at $x = 5.292$

g. State your conclusion with a complete sentence.

Since our f statistic computed: $8.836 > 5.2922$, there is strong evidence to reject H_0. Thus, at least two of the means are not the same.

Score: /7