Stat 5021 Midterm Exam, Spring '15 Name ____________________

This test was taken by 39 people. Each part was worth 10 points for a total possible score of 100. There were 2 scores in the 90's with the high score being 97. There were 4 in the 80's, 13 in the 70's, 8 in the 60's, 6 in the 50's, 3 in the 40's and 2 in the 30's with a low score of 5.

You may use one 8" by 11" formula sheet (both sides) but you **may not use** any electronic computing device. There is no need to reduce numerical formulas to their simplest form. Your answers may contain R commands.

1. Let the joint distribution of X and Y be given in the table below.

		Y			
		1	2	3	4
X	0	.15	.10	.10	.20
	1	.05	.25	.10	.05

i) Find the probability distribution of Y.

ii) Find the expected value of Y.

iii) Find the covariance of X and Y.

Solution: i) Y takes on the values 1, 2, 3 and 4 with probabilities .20, .35, .20 and .25.

ii)

$$E(Y) = \sum_y yP(Y = y) = 1 \times .20 + 2 \times .35 + 3 \times .20 + 4 \times .25 = 2.5$$

iii) Now $Cov(X, Y) = E(XY) - E(X)E(Y)$ and since $E(X) = .45$ we need to find

$$Cov(X, Y) = \sum_{(x,y)} xyP(X = x, Y = y) = 1\times.050+2\times.25+3\times.10+4\times.05 = 1.05$$

$\blacksquare$

2. Let $X_1, \ldots, X_{50}$ be independent identically distributed normal random variables where $E(X_i) = 70$ and $V(X_i) = 16$

i) Does the random variable $(X_1 + X_2)/2$ have a smaller standard deviation than the random variable $(X_3 + X_4 + X_5)/3$? Explain.

ii) For each i let

$$\begin{aligned} Y_i &= 1 \quad \text{if} \quad X_i > 70 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

Let $\bar{Y} = (Y_1 + Y_2 + \cdots Y_{50})/50$. Find $E(\bar{Y})$ and $V(\bar{Y})$.

Solution: i) $V((X_1 + X_2)/2) = (1/4)(16 + 16) = 8$ while the variance of the second is $16/3$.

ii) Note $50 * \bar{Y}$ has a binomial distribution with $n = 50$ and $p = 1/2$ so $E(\bar{Y}) = 1/2$ and $V(\bar{Y}) = 0.25/50$

∎

3. To compare two different types of automobile shock absorbers say A and B, four shocks of each type were purchased. Then one of each type was randomly assigned to the rear wheels of four automobiles. After the cars had been driven 20,000 miles the strength of each shock was measured, coded and recorded. Using the results given in the table below, find a 95% confidence interval for the true average difference between the strength of shock A and shock B.

car	1	2	3	4
Shock A	28.8	22.5	24.6	29.5
Shock B	28.4	22.1	24.0	29.3

Solution: This is paired data since each of the four cars gets one shock of each type. Let μ_A and μ_B be the true average strengths of the two shocks. Let $X_{A,i}$ and $X_{B,i}$ be the observations from the ith car. Then the $x_{A,i} - x_{B,i}$'s for this data are 0.4, 0.4, 0.6 and 0.2. Now using the formula for paired data and the fact that $qt(0.975) = 3.18$ we have

$$0.4 \pm \frac{3.18}{\sqrt{4}} \sqrt{0.08/3}$$

Another acceptable solution using R commands would be:

```
A<-c(28.8,22.5,24.6,29.5)
B<-c(28.4,22.1,24.0,29.3)
t.test(A,B,paired=T)
```

∎

4. Suppose at the University of Minnesota 75 of 110 freshman from the state of Minnesota passed a language placement exam while 55 of 100 freshman from outside the state of Minnesota passed the exam. Let p_m denote the true proportion of Minnesota students who would pass the exam and p_0 denote the true proportion of non-Minnesota residents who would pass the exam.

i) Write a mathematical expression for a 99% confidence interval for $p_m - p_o$.

ii) For testing $H : p_m = p_o$ against $K : p_m \neq p_o$ find the p-value or level of significance for these data.

Solution:

i)For the first part I wanted a formula and gave only partial credit to answers using prop.test.

$$\frac{75}{110} - \frac{55}{100} \mp qnorm(0.995) \sqrt{\left(\frac{75}{110}\frac{35}{110}\right)/110 + \left(\frac{55}{100}\frac{45}{100}\right)/100}$$

ii) For this part you can get the p-value of 0.0684 using

```
prop.test(c(75,55),c(110,100))$p.value
```

∎

5. Suppose that 1% of batteries produced by a factory are truly defective. There is test to check to see if a battery is defective. The test performs as follows.

- Among batteries that are actually defective 98% will be labeled defective by the test.
- Among batteries that are not defective 3% will be labeled defective by the test.

i) Find the probability that a battery selected at random from the factory will be labeled not defective.

ii) Given that it is labeled not defective find the probability that it is actually not defective.

Solution:

i)

$$P(LND) = P(ND)P(LND \mid ND) + P(D)P(LND \mid D)$$
$$\frac{99}{100}\frac{97}{100} + \frac{1}{100}\frac{2}{100}$$

ii)

$$P(ND \mid LND) = P(ND \cap LND)/P(LND) = (99 \times 97)/(99 \times 97 + 1 \times 2)$$

∎