1. (32 points) A person stands in a shopping mall. She stops 400 people and asks each how many times per month they go out for hamburgers, and which of two national hamburger chains they prefer. For the 400, the average number of burger runs per month is 20, with an SD of 15, and 250 say they prefer MacTavish’s hamburgers to Burger Empire’s.

(a) Assuming that she has managed to get a simple random sample of people in the city, what is a 95% confidence interval for the average number of visits to a hamburger restaurant for a person in that city?

(b) How many people would she have had to question to reduce the size of her confidence interval by a factor of 3 (assuming a similar average and SD)?

(c) Burger Empire claims that 45% of the people in the city prefer their burgers to MacTavish’s. Is the surveyor’s result (statistically) significantly different from Burger Empire’s claim?

(d) Describe one possible source of error or bias in the data.

2. (30 points) Two hundred people have applied for the open position in Colgate’s Math Dept. Colgate’s advertisement expressed some preference for certain fields of specialization, but Prof. Tucker doubts that the ad made any difference to the people who applied — if someone was interested in a job at Colgate, he/she applied whether or not the ad said we wanted someone in that field. Here are the numbers in each field who applied, and the estimated percentages of all math Ph.D.’s who specialize in each area:

<table>
<thead>
<tr>
<th>Field</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>90</td>
<td>40%</td>
</tr>
<tr>
<td>Applied Math</td>
<td>30</td>
<td>25%</td>
</tr>
<tr>
<td>Discrete Math</td>
<td>20</td>
<td>10%</td>
</tr>
<tr>
<td>Topology/Algebra</td>
<td>60</td>
<td>25%</td>
</tr>
</tbody>
</table>

(a) If Prof. Tucker is correct, how many of each specialty would have applied?

(b) Which test of significance should he use to decide whether he was correct? Compute the relevant values (except for the $p$-value — save that for part (c)).

(c) Should Prof. Tucker conclude that the ad’s expression of preference made no difference? Explain.

3. (16 points) A “big die” has 10 (pentagonal) faces, numbered 1 through 10, all equally likely to come up if it is rolled.

(a) What is the probability of getting a sum of exactly 17 if a big die is rolled twice?

(b) On one roll of a big die, are the events “getting at most 4” and “getting an odd number” independent?

(c) What is the probability of getting at least one 3 on 4 rolls of a big die?

(d) Which is probably closer to correct, to use normal approximation to estimate the result of the 50th roll of a big die (before it is actually rolled), or to estimate the total of 50 rolls of a big die? Why?
4. *(15 points)* The local PTA runs a monthly lottery to raise money (thereby setting a terrible example for children). It sells tickets for a dollar, for a single prize of $500 (so the winner is ahead $499 — he/she doesn’t get the dollar back), and its 600 tickets sell out every month. The tickets are numbered, so there are no ties.

(a) You buy one ticket every month. What should you expect your winnings to total in 10 years (120 months), give or take how much?

(b) You and 9 friends buy 10 tickets, agreeing to split the pot 10 ways if any one of your tickets wins. Same questions as in (a). (Note: If your group wins, you are ahead only $49 — you don’t specifically get back the dollar you paid.)

(c) With which scheme, (a) or (b), are you more likely to break even?

5. *(20 points)* A (misguided) day trader tries to use linear regression to project the values of stocks from one day to the next. He follows 100 stocks and finds that their prices yesterday averaged 240, with an SD of 100, and that their prices today averaged 250 with an SD of 200. The correlation coefficient was 0.9.

(a) For a stock that cost 280 yesterday, what would he project is its value today?

(b) How far should we expect his projection to be off?

6. *(12 points)* Concerning the article ”The face of poverty: Is it what you think?” by Greg Jonsson: Do you find the statistics presented in the article helpful? What do they imply (if anything) about “solutions to poverty” – do they suggest that monetary support, or education, or job creation, or ... (or even nothing) would be most helpful?
Math 102 / Core 143 AX/BX— Solutions to Final Exam I

1. (a) 20 ± 2(15/\sqrt{400}) = 20 ± 1.5 visits per month.
   (b) To reduce the size of the interval to 20 ± 0.5 visits per month, she would have had to interview 9 times as many people, or 3600.
   (c) Her result was 250 out of 400 prefer “Mickey T’s”, more than the 220 that BE’s figure would give. So if BE’s claim were correct (null hypothesis), the SE of the count would be 
   \((1 - 0)\sqrt{(0.55)(0.45)}/\sqrt{400} \approx 9.9\), so the probability that she would have gotten her result was 
   \[P(count \geq 250) = P(z \geq \frac{250 - 220}{9.9} \approx 3) < 5\%\ ;\]
   so her result is significantly different.
   (d) For one thing, she was interviewing in a shopping mall, so her subjects were probably more likely to eat out than the average person.

2. (a) Analysis 80, Applied Math 50, Discrete Math 20, Topology/Analysis 50.
   (b) A \(\chi^2\)-test, with 4 - 1 = 3 degrees of freedom:
   \[\chi^2 = \frac{(90 - 80)^2}{80} + \frac{(30 - 50)^2}{50} + \frac{(20 - 20)^2}{20} + \frac{(60 - 50)^2}{50} = 11.25\ .\]
   (c) From the \(\chi^2\)-table, with 3 degrees of freedom, the probability of getting a \(\chi^2\)-value of at least 11.25 is less than 5%, so Prof. Tucker should conclude that the ad made some difference (probably in encouraging people in Applied Math to apply).

3. (a) There are 4 ways to get a 17 out of 100 possible rolls, so the probability is 4%
   (b) The probability of getting an odd number is 50%. There are 4 ways to get at most 4, of which 2 are odd, so the conditional probability of getting an odd number given that the number is at most 4 is 2/4, equal to 50%. So the events are independent.
   (c) One minus the probability of never getting a 3 on 4 rolls, i.e., \(1 - .9^4 = .3439\)
   (d) The distribution of a single roll (whether the first or the 50th) is uniform, not normal, so normal approximation is not a good way to estimate the result of a single roll. But the distribution of totals of 50 rolls is probably close to a normal distribution, so normal approximation is reasonable.

4. (a) The EV of the sum is \(120\left(\frac{1}{500}(499) + \frac{599}{600}(-1)\right) = \frac{499 - 599}{5} = -20\) dollars, with an SE of \(\sqrt{120}(499 - (-1))\sqrt{\left(\frac{1}{500}\right)\left(\frac{599}{600}\right)} = \sqrt{120}(500)\sqrt{599} \approx 223\) dollars.
   (b) The EV of the sum is \(120\left(\frac{10}{500}(49) + \frac{590}{600}(-1)\right) = \frac{490 - 590}{5} = -20\) dollars, with an SE of \(\sqrt{120}(49 - (-1))\sqrt{\left(\frac{10}{500}\right)\left(\frac{590}{600}\right)} = \sqrt{120}(50)\sqrt{590} \approx 70\) dollars.
   (c) Since the EV’s of the schemes are the same negative value, and the SE is higher with the first, you are more likely to get at least 0 winnings with the first scheme.

5. (a) 280 is \((280 - 240)/100 = .4\) in standard units, so we should project today’s price in standard units as \(.9(0.4) = .36\), or in “real” units (dollars?), \(.36)(200) + 250 = 322\.
   (b) \(200\sqrt{1 - (.9)^2} \approx 87\).

6. The question asks mostly for your reaction. I will be looking for evidence that you have read the article; for example, a comment that a larger portion of the population than might have been expected have experienced poverty, or a reference to the “musical chairs” analogy.