Instructions. To get full credit, you must show your work. Good luck!

1. (18 pts.) True or false:
   (a) If one of the $y$ values is negative, the correlation coefficient between $x$ and $y$ will be negative.
   (b) Quota sampling is a probability method for sampling.
   (c) If the $P$-value is 0.9%, there is less than a 1% chance for null hypothesis to be right.
   (d) Big values of $P$ are bad for the null hypothesis.
   (e) If a coin is tossed 1,000 times, then the chance of getting between 490 and 510 heads is lower than the chance of getting between 48% and 52% heads.
   (f) If six hundred draws are made at random with replacement from the box $[0, 0, 1]$, then the expected value for the percentage of 1’s among the draws will be around $33\frac{1}{2}$%, give or take 2% or so.
   (g) The SD of the box $[1, 0, 0, 0, 0, 1]$ is the same as the SD of the box $[1, 0, 0, 0]$.
   (h) With a well-designed sample survey, the sample percentage is very likely to equal the population percentage.
   (i) A 99.7%-confidence interval for the population average is the interval “sample average ± 3SEs”.

2. (9 pts.) A coin will be tossed 10 times. What is the chance of obtaining
   (a) exactly 1 head?
   (b) no heads?
   (c) at least two heads?

3. (12 pts.) The town of Hayward (California) has about 50,000 registered voters. A political scientist takes a simple random sample of 500 of these voters. In the sample, the breakdown by party affiliation is
   
   Republican 115  
   Democrat 331  
   Independent 54

   (a) Among all registered voters in Hayward, the percentage of democrats is estimated as _____.
   (b) This estimate is likely to be off by _____ or so.
   (c) The range from _____ to _____ is a 95%-confidence interval for the percentage of democrats _____.

   Fill in the blanks. (The first four blanks are filled in with numbers; the last blank takes a phrase—10 words or less.)

4. (8 pts.) Before the strike of 1994, the median salary of the 746 major league baseball players was about $500,000. The lowest salary was about $100,000 and the highest was over $5,000,000. Choose one option and explain:
   (i) The owners were paying out substantially less than $373 million per year to the players.
   (ii) The owners were paying out substantially more than $373 million per year to the players.
5. (14 pts.) A certain town has 25,000 families. These families own 1.6 cars, on the average; the SD is 0.90. And 10% of them have no cars at all. As part of an opinion survey, a simple random sample of 1,500 families is chosen. To estimate the chance that the average number of cars owned by the sample families is greater than two, a box model is needed.

(a) Should the number of tickets in the box be 25,000 or 1,500?
(b) Each ticket in the box shows

\begin{align*}
\text{a zero or a one} & \quad \text{number of cars owned}
\end{align*}

(c) True or false: the SD of the box is 0.90.
(d) True or false: the number of draws is 1,500.
(e) Find the chance (approximately) that the average number of cars owned by the sample families is greater than two.

6. (8 pts.) In any survey, a fair number of people who are in the original sample cannot be contacted by the survey organization, or are contacted but refuse to answer questions. A high non-response rate is a serious problem for survey organizations. True or false, and explain: this problem is serious because the investigators have to spend more time and money getting additional people to bring the sample back up to its planned size.

7. (8 pts.) Four hundred draws are made at random with replacement from a box of numbered tickets. If the average of the draws is 20 and the SD of the draws is 20, find a 99.7%-confidence interval for the average of the box.

8. (8 pts.) The following tables show the distributions of gender and classification in two statistics classes with 60 students each.

<table>
<thead>
<tr>
<th>Math 1680-003</th>
<th>Men</th>
<th>Women</th>
<th>Math 1680-007</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshmen</td>
<td>20</td>
<td>10</td>
<td>Freshmen</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Sophomores</td>
<td>10</td>
<td>20</td>
<td>Sophomores</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td></td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td></td>
<td>20</td>
<td>40</td>
</tr>
</tbody>
</table>

(a) Say whether gender and classification are dependent or independent in each class.
(b) True or false, and explain: In section 003, there are 50% men and 50% freshmen, so the chance of picking a male freshman by random selection is $\frac{1}{2} \times \frac{1}{2}$.
(c) True or false, and explain: In section 007, there are 33$\frac{1}{3}$% men and 50% freshmen, so the chance of picking a male freshman by random selection is $\frac{1}{3} \times \frac{1}{2}$.
(d) True or false, and explain: In section 007, the chance of picking a sophomore or a man by random selection is $\frac{1}{2} + \frac{1}{4}$.

9. (15 pts.) One hundred investigators set out to test the null hypothesis that a die shows a $\blacksquare$ 16$\frac{2}{3}$% of the time. Each investigator rolls the die 600 times, counts the number of $\blacksquare$’s, and does a $z$-test. The results are plotted below.
Investigator #93 gets 71 $\blacksquare$’s. Fill in the blanks for this investigator.

(a) This investigator gets a $\blacksquare$ ___ % of the time.

(b) To make a box model for the null hypothesis, we need one ticket in the box for each _________ face of the die roll in the sample.

(c) The ticket is marked ___ for $\blacksquare$ and ___ otherwise.

(d) The number of tickets in the box is ___ and the number of draws is ___.

(e) The null hypothesis says that the sample is like ___ ______ made at random with replacement from the box. (The first blank must be filled in with a number; the second, with a word.)

(f) The percentage of 1’s in the box is _______. Options: 16.2%, 11.83%.

(g) The observed number of $\blacksquare$’s is _______.

(h) The expected number of $\blacksquare$’s is _______.

(i) If the null hypothesis is right, the number of $\blacksquare$’s in the sample is like the ______ of the draws from the box. Options: sum, average.

(j) The SE for the number of $\blacksquare$’s is _______.

(k) $z = _____$ and $P = _____$. (The remaining questions of this problem are bonuses: 6 pts)

(l) Does investigator #93 agree with the null hypothesis?

(m) True or false: The $z$-statistic is negative when the number of $\blacksquare$’s is less than 100.

(n) If the null hypothesis is true, how many investigators should get a negative $z$-statistic?

(o) If the null hypothesis is true, how many of them should get a $z$-statistic smaller than $-2$? How many of them actually do?

(p) True or false: The majority of investigators reject the null hypothesis.

10. (Bonus: 8 pts.) Someone is going to play roulette 1000 times, betting a dollar on a column each time. (There are 38 numbers on a roulette table and a column covers 12 numbers. A column bet pays 2 to 1.)

(a) Find the expected value and standard error for the net gain.

(b) Find the chance of coming out ahead. (Coming out ahead means the net gain is positive.)

11. (Bonus: 6 pts.) For entering freshmen at a certain university, scores on the Math SAT and Verbal SAT can be summarized as follows:

- average M-SAT = 525, SD = 125
- average V-SAT = 475, SD = 115, $r = 0.66$

The scatter diagram is football-shaped. One student is chosen at random and has an M-SAT of 600. You would guess his V-SAT is ______ points. Fill in the blank; explain briefly.