

1. (10 pts) The first-semester enrollment at HSC was 1120 students. If all entering freshmen classes were the same size and there were no attrition, then one would expect the 1120 students to be divided equally among freshmen, sophomores, juniors, and seniors. The actual enrollments, by class, are

		Class	Freshman	Sophomore	Junior	Senior
Enrollment	Actual		314	305	261	240
	Expected					

- (a) Fill in the expected enrollments under the above hypothesis.
- (b) Test the hypothesis that enrollment is uniform across the classes. Use $\alpha = 0.05$.
2. (10 pts) Normally when one rolls a die, he shakes it up in his hand and then rolls it. That procedure produces a uniform distribution of numbers 1 through 6. I was wondering whether the distribution would still be uniform if I rolled the die without shaking it up first. So I rolled an ordinary die 120 times without shaking it up. Each time, I began with the number 1 facing up and just rolled the die onto the table. The following table shows the results.

Number	1	2	3	4	5	6
Count	10	14	16	25	23	32

Test the hypothesis that the results come from a uniform distribution, i.e., the hypothesis that $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$, at the 1% level of significance.

3. (10 pts) I modified the `randInt` function on my TI-83 in such a way as to simulate a loaded die that lands 1, 2, or 3 each with probability 10%, 4 or 5 each with probability 20%, and 6 with probability 30%. Then I tested my creation by simulating 100 rolls of the loaded die. The results are shown in the following table.

Value	1	2	3	4	5	6
Frequency	8	11	14	29	15	23

Test the hypothesis that the distribution from which these numbers were taken is 10%, 10%, 10%, 20%, 20%, 30% for the values 1, 2, 3, 4, 5, 6, respectively. Test at the 1% level of significance.

4. (12 pts) The Richmond Times-Dispatch reported the number of incidents of crime for 24 cities and counties in the central Virginia area for 2007 and 2008. The following table shows the data for a random sample of 10 of those cities and counties.

City/County	No. of Incidents	
	2007	2008
Ashland	312	283
Colonial Heights	799	852
Dinwiddie County	485	553
Goochland County	223	234
Hanover County	1140	1189
Henrico County	10518	10795
Hopewell	1166	1302
King William Co.	58	50
Petersburg	2947	3154
Richmond	10971	9759

- (a) (10 pts) Use the count data to test the hypothesis that the two distributions are the same from 2007 to 2008, that is, the populations are homogeneous. Test at the 5% level of significance.
- (b) (2 pts) Under the assumption of homogeneity, find the expected number of incidents of crime in Ashland in 2007.
5. (12 pts) According to an article in the *Richmond Times-Dispatch*, Nov. 14, 2008, there were 6190 vehicle crashes involving deer in Virginia last year.¹ The article divided the crashes into three categories depending on the type of road: interstate highway, primary road, or secondary road. One might expect the crashes on the interstate to be the most severe and the ones on the secondary roads to be the least severe because of the difference in speed. One might be wrong. The following table shows the number of accidents and the number of injuries in each category.

Type of road	Interstate	Primary	Secondary
No. of crashes	765	3505	1920
No. of injuries	59	281	145

- (a) (10 pts) Test the hypothesis that the two distributions are the same, that is, the populations are homogeneous. Test at the 5% level of significance.
- (b) (2 pts) Find the expected number of injuries for each category under the assumption of homogeneity.
6. (12 pts) Every year the National Survey of Student Engagement (NSSE) is conducted on all participating campuses. HSC has participated in 2003 and 2007. One of the questions on the 2003 survey was “Discussed ideas from your readings or classes with others outside of class” and the student responds with “Never,” “Sometimes,” “Often,” or “Very Often.” The table below shows the results for HSC and for all liberal arts colleges.

¹<http://www.inrich.com/cva/ric/search.apx.-content-articles-RTD-2008-11-14-0161.html>

	Never	Sometimes	Often	Very Often
HSC	7	54	52	39
All LA colleges	266	2305	3142	2468

- (a) (8 pts) Conduct a test, at the 5% level of significance, to determine whether the two populations (HSC students, all LA college students) are homogeneous.
- (b) (2 pts) According to the null hypothesis, what is the expected number of HSC students who very often discussed ideas from reading with others outside of class?
7. (12 pts) In a sample of 42398 traffic deaths in 2002, investigators determined whether speeding was a factor and whether the road was straight or curved at the site of the accident. They reported the following data.

	Straight	Curved
Speeding	8359	5304
Not Speeding	23139	5596

- (a) (2 pts) Identify the two variables in this study.
- (b) (8 pts) Conduct a test, at the 5% level of significance, to determine whether the two variables are independent.
- (c) (2 pts) According to the null hypothesis, what is the expected number of traffic deaths on straight roads in which speeding was a factor?
8. (15 pts) I was looking at the birthdays of my Facebook friends, both of them. No, actually I have many of Facebook friends. One would expect their birthdays to be evenly distributed over the 12 months of the year, because I do not choose my friends based on their birthdays. So I checked and obtained the following data.

Year	Jan - Mar	Apr - Jun	Jul - Sep	Oct - Dec
No. of Birthdays	22	14	19	15

- (a) (3 pts) Fill in the expected number of birthdays for each three-month period under the assumption that birthdays are evenly distributed over the year.
- (b) (12 pts) Test the hypothesis at the 10% level of significance that birthdays are evenly distributed over the year.
9. (17 pts) An article in the Richmond-Times Dispatch, titled “The word is out: police are watching boaters²” reports the following data for number of boating accidents and number of deaths resulting from those accidents.

²<http://www.reuters.com/article/2011/10/11/us-prostate-cancer-idUSTRE79A6VE20111011>

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
No. of Accidents	169	152	122	152	132	137	145	95	136	105
No. of Deaths	19	20	20	20	12	23	23	17	25	15

A natural question is to ask whether the number of deaths per year exhibits the same distribution as the number of accidents.

- (a) (12 pts) Test the hypothesis at the 5% level of significance that the two variables are independent.
- (b) (3 pts) If they do represent the same distribution, then what are the expected number of accidents and the expected number of deaths in 2001? Fill in the expected counts for both rows in the first two columns (2001 and 2002).
- (c) (2 pts) How many degrees of freedom are there in this test?