Applied Statistics Comprehensive Examination

Statistical Methods I & II

You must give complete explanations to receive full credit.

1. The data in the table below are the counts of at bats (attempts) and hits (successes) for Philadelphia Phillies first baseman Ryan Howard over the 2009 to 2012 seasons.

Year	At Bats	Hits
2009	616	172
2010	550	152
2011	557	141
2012	260	57

(a) (20 points) Using $\alpha = 0.05$, test the theory that Howard's success rate is constant over the four years.

(b) (5 points) Estimate the p-value for the test in (a) by specifying an interval where it must fall.

2. Large shipments of a certain item are received at a factory. When each shipment arrives, a quality control worker selects ten items at random and inspects the selected items. If more than one of the ten items is defective, then the entire shipment is rejected. Otherwise, the entire shipment is accepted. Suppose that we think of this procedure as a test of $H_0 : \pi = 0.02$ against $H_a : \pi > 0.02$, where π is the proportion of defective items in the shipment.

(a) (10 points) What is the α level for this test?

(b) (10 points) What is the power of the test when $\pi = 0.2$?

3. In a drug experiment involving 12 patients suffering from a particular skin condition, six patients were assigned to each of two drugs. One outcome variable was the number of days until recovery from the condition. Data for this outcome variable appear below.

Drug	Recovery Times (days)						$\bar{x}$	s
A	13	6	9	9	10	9	9.33	2.25
B	10	13	10	19	12	10	12.33	3.50

(a) (15 points) Using $\alpha = 0.05$, test for a difference in the variance of the number of days to recovery for the two drugs.

(b) (10 points) Find and interpret a 95% confidence interval for the difference in mean recovery time for the two drugs.

4. An experiment was run to compare the average yields for three varieties of wheat. Four fields were randomly assigned to each of the three varieties, and the yields in bushels per acre are given in the table below.

Variety	Yields				$\bar{x}$	s
A	38	40	37	39	38.50	1.29
B	36	36	39	37	37.00	1.41
C	53	48	43	37	45.25	6.85

(a) (15 points) Give a complete ANOVA table, and test equality of means at level 0.05.

(b) (10 points) Apply Fisher's protected LSD. Which pairs of means are declared significantly different?

(c) (5 points) List the assumptions needed for the procedures that you applied in parts (a) and (b). Without doing any formal testing, discuss whether any of the assumptions seem to be violated or not.