Solutions to STAT 300 Midterm Exam (Version I)

Problem 1.

1. T 2. F 3. T 4. F 5. T

6. T 7. T 8. F 9. F 10. F.

Problem 2.

1. A 2. A. 3. D. 4. C. 5. B

6. C 7. B. 8. D. 9. D. 10. A.

Problem 3.

(1) (2 pts each) (i) A (ii) C

(2) (2 pts each) (i) 6.1 (ii) 0.49 (or 0.50) (iii) 12.2 (iv) $F(2,\ 57)$ or $F_{2,57}$ (v) 0.49 (or 0.50)

(3) (2 pts each) 0.0167, because we are doing three tests simultaneously.

(4) (2 pts each) Method I: paired t-test, Method II: sign test. The reason is that data for each subject in this design are paired (or correlated).

(5) (1 pt, 2 pts) Interaction. When there is an interaction, we should analyze data for male and female separately.

(6) $H_0 : \frac{\mu_1+\mu_2}{2} = \mu_3$.

Problem 4.

The data in this study can be summarized by the following 2 by 2 table

	Lab I	Lab II	Total
Pass	17	15	32
Fail	3	5	8
	20	20	40

(1) (2 pts each) (i) Fisher's exact test, χ^2 test. (ii) The Fisher's exact test may be more reliable here since the cell count 3 (and 5) are small (or the sample size is small) .

(2) (2 pts each) $e_{11} = 40 * (32/40) * (20/40) = 16$ and $o_{11} - e_{11} = 17 - 16 = 1$.

(3) No (since there is no randomization).

(4) Let p be the proportion of students who like activity-based labs. The hypotheses to be tested are

$$H_0 : p \leq 0.05 \qquad vs \qquad H_1 : p > 0.5.$$

Let X be the number of students in Lab I who like activity-based labs. Then, under H_0, we have

$$X \sim Binomial(20, 0.5).$$

(1) (4 pts) The significance level is

$$\alpha = P(reject\ H_0|H_0\ holds) = P(X \geq 14|H_0\ holds) = \sum_{k=14}^{20} C_{20}^k 0.5^k \times 0.5^{20-k}.$$

(2) (4 pts) The power is

$$power = P(reject\ H_0|H_1\ holds) = P(X \geq 14|p = 0.6) = \sum_{k=14}^{20} C_{20}^k 0.6^k \times 0.4^{20-k}.$$

(3) (2 pts) Note that, under H_0, $E(X) = np = 20 \times 0.5 = 10$ and $Var(X) = np(1-p) = 5$. Thus, the normal approximation for the null distribution of X is $N(10,\ 5)$ (or $N(0.5, 0.25)$ if one use $\hat{p} = X/n$ as a test statistic). (Note that here we use $N(\mu, \sigma^2)$ to denote a normal distribution. Some may also use notation $N(\mu, \sigma)$.)