Due by 4pm on Friday, February 21. Please leave your homework on the table before class begins on Friday or leave it in the dropbox outside my office. Do not forget to attach the honor code. There are five problems, 20 points each.

(1) (a) Use the pchisq() function in R with $v = 1$ df to find the following p-values. Round to 4 decimal places.

(i) $p = P(\chi^2 > 3.8416)$ (iii) $p = P(\chi^2 > 1)$ (v) $p = P(\chi^2 > 0)$

(ii) $p = P(\chi^2 > .5625)$ (iv) $p = P(\chi^2 > 2.706)$

(b) Use the qchisq() function in R to compute the following (right-tail) critical values. Round to 4 decimal places.

(i) $\chi^2_{0.05,\nu=4}$ (iii) $\chi^2_{0.01,\nu=6}$ (v) $\chi^2_{0.1,\nu=1}$

(ii) $\chi^2_{0.025,\nu=4}$ (iv) $\chi^2_{0.005,\nu=4}$

(2) A coin is flipped 1,200 times and 732 heads are observed. The coin flipper wants to know whether or not the coin is fair. Use a Wald test with the χ^2 distribution and a 2% level of significance to determine if there is evidence that the coin is not fair. Make sure to state the relevant hypothesis, calculate either a p-value or rejection region, and state a conclusion in the context of the problem.

(3) To collect data in an introductory statistics course, recently I gave the students a questionnaire. One question asked whether the student was a vegetarian. Of 25 students, 0 answered "yes." They were not a random sample, but let us use these data to illustrate inference for a proportion. Let p denote the population proportion who would say "yes." Consider $H_0 : p = 0.50$ and $H_a : p \neq 0.50$.

(a) What happens when you try to conduct the "Wald test," for which $z = \dfrac{(p - p_0)}{\sqrt{\frac{p(1-p)}{n}}}$ uses the estimated standard error?

(b) Find the 95% "Wald confidence interval" for p. Is it believable? (When the observation falls at the boundary of the sample space, often Wald methods do not provide sensible answers.)

(c) Conduct the "score test," for which $z = \dfrac{(p - p_0)}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ uses the *null* standard error. Report the p-value.

(d) Verify that the 95% score confidence interval (i.e., the set of p_0 for which $|z| < 1.96$ in the score test) equals $(-0.196, 0.196)$.

(4) Let $y_1, y_2, \ldots, y_n$ be a random sample from a normal pdf with unknown mean μ and variance 1. Find the form of the likelihood ratio test statistic for $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$.

(5) Two different companies have applied to provide cable television service in a certain region. Let p denote the proportion of all potential subscribers who favor the first company over the second. Consider testing $H_0 : p = 0.5$ versus $H_a : p \neq 0.5$ based on a random sample of 25 individuals. Let the RV X denote the number in the sample who favor the first company and x represent the observed value of X.

(a) Which of the following rejection regions is most appropriate and why?

$$R_1 = \{x : x \leq 7 \,\text{or}\, x \geq 18\} \qquad R_2 = \{x : x \leq 8\} \qquad R_3 = \{x : x \geq 17\}$$

(b) In the context of this problem situation, describe what Type I and Type II errors are.

(c) What is the probability distribution of the test statistic X when H_0 is true? Use R to compute the probability of a Type I error.

(d) Use R to compute the probability of a Type II error for the selected region when $p = 0.3$.