

STATISTICS COMPREHENSIVE EXAM
JANUARY 2014

Instructions: No references are permitted during the exam. Solutions must be written legibly and neatly on separate sheet(s) of paper with your name and problem number at the top of each page. Be sure to provide complete and clear reasons for all of your steps (e.g. no statement is “clear” or “obvious” unless it’s a definition or given as an assumption, and do not only name a theorem. Also show that its hypotheses are satisfied.) If you are asked to prove a known theorem, do not merely quote that theorem as your proof; instead, produce an independent proof. All problems are equally weighted. You have three (3) hours to submit your solutions.

- (1) Let X_1, \dots, X_n be a random sample from the density

$$g(x) = \frac{\theta}{x^2}, \quad 0 < \theta \leq x.$$

- (a) Define a random variable $W = \frac{1}{\theta} \ln(X)$. Derive the distribution of W .
(b) Find a sufficient statistic for θ per the distribution of W .

- (2) Let w_1, \dots, w_n be a sample of data from the distribution of W in Problem (1).

- (a) Let θ have an exponential distribution, namely $\theta \sim \lambda e^{-\lambda\theta}$. What is the posterior distribution of θ , given a data set w_1, \dots, w_n ?
(b) How much data should be used and what should the value of λ be so that the mean and variance of the posterior distribution are 1.5 and 0.15, respectively? Let $\bar{w} = 1$.

- (3) Let X_1, \dots, X_{20} be i.i.d. from a distribution having a p.d.f. of the form

$$f(x) = \theta e^{-\theta x}, \quad x \geq 0.$$

Find the Rejection Region of the most powerful test for $H_0 : \theta = 1$ versus $H_1 : \theta = 2$. Let $\alpha = 0.05$. HINT: Find the rejection region in terms of x (i.e. $g(x) < c^*$). Then determine the distribution of $g(x)$. Finally, use a normal approximation to solve for c .

- (4) Approximate the power of the test in Problem (3). Is it a good hypothesis test? Explain.

- (5) Describe the logic behind the Binomial test. That is, state the test statistic. Report the distribution of the test statistic and explain why it has that particular distribution.

- (6) The 2-sample permutation test for independent samples permutes the combined data instead of permuting within each sample. Explain why. That is, what is the rationale behind assigning some values from sample 1 to sample 2 and vice versa?