STAT4106 Exam 2 Study Guide

John W Smith Jr

This is the study guide for the Final Exam. All problems on here have the solutions provided as well as hints to help to get you there. This assignment can be handed in at any time before the final and will be graded based on completion. This is meant to help you practice. You are free to work on these problems with your classmates, though everyone must turn in their own copy of the assignment.

Problem 1: Exponential Sampling Distributions

Suppose that $X_1, .., X_n \sim Exp(\lambda)$, with pdf

$$f(x|\lambda) = \lambda \exp(-\lambda x), x \in (0, \infty)$$

Part A

Find the distribution of the k^{th} order statistic, $X_{(k)}$.

Part B

Find $M_X(t)$, the moment generating function for an $Exp(\lambda)$ distribution.

Part C

Suppose we are interested in making closed form inference regarding the mean of our sample distribution. Find the sampling distribution of $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Problem 2: Just a Normal Problem

Suppose that $X_1, ..., X_n \sim$ iid $Normal(\mu, \sigma^2)$.

Part A

Suppose that we are interested in finding the distribution of a linear combination of random variables, $\sum_{i=1}^{n} a_i X_i$, where $a_i \in \mathbb{R}$. Using the Normal MGF, find the distribution of this linear combination. You do not need to derive the MGF.

Part B

Suppose that we are interested in estimating the variance for the normal distribution by using the estimator

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Find the bias and variance of S^2. Hint: you may want to first find the distribution of $(n-1)S^2/\sigma^2$.

Part C

Show that S^2 is a pivotal quantity for σ^2

Part D

Suppose that we are interested in finding an estimator of the form $\hat{\sigma}^2 = cS^2$. Find the value for c that minimizes the MSE of the estimator.

Part E

Can we construct a confidence interval of the form $cS^2 \pm k\sigma_{cS^2}$? Justify your answer.

Problem 3: Comparing Estimators and Exponential Families

Suppose that X is a random variable belonging to the exponential family, with parameterization

$$f(x|\Theta) = c(\Theta)h(x)\exp\Big(\sum_{i=1}^{n} t_i(x)w_i(\Theta)\Big)$$

Part A

Using the factorization theorem, show that if we observe $X_1, ..., X_n$, then

$$T(\mathbf{X}) = \Big(\sum_{i=1}^{n} t_1(X_i), ..., \sum_{i=1}^{n} t_k(X_i)\Big)$$

is a sufficient statistic for Θ.

Part B

Show that the exponential distribution is a member of the exponential family. Let the parameterization be given by:

$$f(x|\lambda) = \lambda^{-1}\exp(-\lambda^{-1}x)$$

Part C

Using the result from Part A, find a sufficient statistic for an iid sample of Exponential random variables. Is this a consistent estimator? Justify your answer.

Part D

State the Rao-Blackwell Theorem. Using the sufficient statistic obtained above, find the MVUE for Exponential distribution.

Part E

Suppose that instead we parameterize the exponential as

$$f(x|\theta) = \theta\exp(-\theta x), x \in (0, \infty), \theta > 0$$

If we were to use $\hat{\theta} = \frac{1}{\hat{\lambda}_{MVUE}}$, would our $\hat{\theta}$ converge in probability to θ? List any exceptions and justify your answer.

Problem 4: Binomial Bayes Bounds and more

Suppose that we observe $X_1, ..., X_n \sim$ iid $Bern(p)$.

Part A

Find the maximum likelihood estimator for p.

Part B

Find the method of moments estimator for p.

Part C

Compute the Cramer-Rao Lower Bound for p. Do either of the estimators above attain the bound?

Part D

Suppose that we want to perform a Bayesian analysis with $\pi(p) \sim Beta(.5, .5)$. Compute the posterior $\pi(p|\mathbf{X})$. Using this posterior, find a Bayes estimator for p.

Part E

Compute the MSE for the MLE and Bayes estimators. Find respective intervals where each estimator performs better.

Problem 5: Power and LRTs

Part A

Suppose that $X_1, .., X_n \sim N(\mu, \sigma^2)$ with σ^2 known. Let $H_0 : \mu = \mu_0$ vs $H_a : \theta \neq \theta_0$. Show that the test based on the likelihood ratio test statistic is equivalent to the Normal two-sided z-test.

Part B

Now suppose that σ^2 is not known. Show that the test based on the likelihood ratio test statistic is equivalent to the two-sided t-test.

Part C

Suppose that a factory foreman is interested in testing whether or not one of his pieces of machinery needs to be calibrated. To do this, he looks at $n = 50$ items $(X_1, ..., X_n)$ and tests whether or not they are defective. He picks $H_0 : p = .005$ vs $H_a : p > .005$.

Suppose that the foreman decides to use $RR = \{x | \sum_{i=1}^{n} x \geq 2\}$. Compute α, the Type I error rate.

Part D

Suppose that the true value for p is $p = .01$. Compute β, the Type II error rate, as well as the power of the test.

Problem 6: Inverse CDF and Monte Carlo

Suppose that $X \sim Cauchy(0, 1)$, with pdf given by

$$f(x) = \frac{1}{\pi(1 + x^2)}, x \in \mathbb{R}$$

Part A

Find the CDF, $F(x)$.

Part B

Find the inverse cdf, $F^{-1}(x)$

Part C

Write psuedo-code to outline how to generate a random sample from the pdf $f(x)$. Are these exact samples?

Part D

Suppose that you have mastered the art of darts and have decided to give yourself a challenge by playing on an elliptical dartboard, $.5x^2 + y^2 = 1$. You estimate that your dart-throwing skills in the x and y dimension roughly follow a $Beta(4, 4)$ distribution. Outline a Monte Carlo algorithm to estimate the probability that you hit the board.