BASKIN SCHOOL OF ENGINEERING

SAM Program, Statistics Track

2018 First Year Exam: June 11, 2018

INSTRUCTIONS

Please complete all problems on the **supplied exam papers**. Write your exam ID number and problem number on each page. Use only the **front side** of each page.

Some formulas that may be useful:

- The density function of the exponential distribution with rate parameter $\lambda > 0$ (mean $1/\lambda$, variance $1/\lambda^2$) is given by
$$f(x) = \lambda \exp(-\lambda x), \quad x > 0 \, .$$

- The density function of the gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$ (mean α/β, variance α/β^2) is given by
$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x > 0 \, .$$

- The density function of the beta distribution with shape parameters $\alpha > 0$ and $\beta > 0$ is given by
$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1 \, .$$

- The density of the generalized inverse Gaussian distribution with parameters $a > 0$, $b > 0$, and $p \in \mathbb{R}$ is given by
$$f(x) = C\, x^{p-1} \exp\{-(ax + b/x)/2\}, \quad x > 0 \, ,$$
where C is the normalizing constant.

- The probability mass function of the Poisson distribution with mean $\mu > 0$ is given by
$$f(x) = \frac{e^{-\mu}\mu^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Problem 1 (AMS 203):

1. (60%) Let $X_1, \ldots, X_n$ be an independent and identically distributed (i.i.d.) sample from an exponential distribution with rate parameter λ. Let $Y = X_1 + \ldots + X_n$.

   (a) (25%) Show that Y follows a gamma distribution with shape parameter n and rate parameter λ.

   (b) (15%) Identify the asymptotic distribution of $\sqrt{n}\{(Y/n) - (1/\lambda)\}$, and specify its parameters. Justify your answer.

   (c) (20%) Consider a random variable Z with conditional distribution, $Z \mid Y = y$, given by a Poisson distribution with mean y. Derive the marginal distribution of Z.

2. (40%) Let X_1, X_2 be i.i.d. exponential random variables with rate parameter λ. Let U be a random variable which is uniformly distributed on $[0, 1]$. Suppose that U, X_1, and X_2 are independent. Let $Z = (X_1 + X_2)U$. Prove that Z is an exponential random variable with rate λ.

Problem 2 (AMS 205B):

Let $X_1, ..., X_n \overset{iid}{\sim} U(-\theta, \theta)$, where $\theta > 0$ is an one-dimensional unknown parameter.

1. (30%) Is $(X_{(1)}, X_{(n)})$ minimal sufficient for θ? Provide an argument supporting your answer.

2. (20%) Let $h(X_1, ..., X_n) = \max\{|X_1|, ..., |X_n|\}$. Prove that $h(X_1, ..., X_n)$ is a complete statistic for θ.

3. (20%) Show that the conditional distribution of $h(X_1, .., X_n)$ given $|X_1|/|X_3|$ is the same as the marginal distribution of $h(X_1, ..., X_n)$.

4. (30%) Provide the likelihood ratio test statistic for testing $H_0 : \theta = 2$ vs. $H_1 : \theta \neq 2$.

Problem 3 (AMS 206B):

We are given a coin and are interested in the probability θ of observing heads when the coin is flipped. An experiment is conducted by flipping the coin (independently) in a series of trials. Suppose we have the observations of x heads and y tails (let $n = x + y$).

1. (25%) Suppose that we do not have enough information to specify a model for the "series of trials" and consider two possibilities:

   • (Binomial) Suppose a fixed number of independent Bernoulli trials, n, are performed and let y be the number of tails out of the n trials.

   • (Negative Binomial) Suppose that independent trials are performed until the y-th tail occurs and we let n denote the number of trials needed for this.

   Will the inference for θ be the same under the binomial and the negative binomial distribution models or not? Provide a full justification for your answer.

2. (75%) Let $x \mid \theta \sim \text{Binom}(n, \theta)$, with known n. Consider a two-component mixture of beta distributions prior for θ, more specifically:

$$\theta \sim g_1(\theta) = \epsilon \, \text{Be}(\alpha_1, \beta_1) + (1 - \epsilon) \, \text{Be}(\alpha_2, \beta_2) \,, \tag{3.1}$$

   with $0 < \epsilon < 1$ and α_1, β_1, α_2, $\beta_2 > 0$.

   (a) (25%) Find the marginal distribution $m_1(x)$.

   (b) (25%) Derive the posterior distribution for θ given x.

   (c) (25%) Consider a Bayesian hypothesis test, $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$, for specified θ_0. Assume the prior under the alternative is $g_1(\theta)$ in (3.1). Let ρ_0 be the prior probability that $\theta = \theta_0$. Find the posterior probability of H_0, $\pi(H_0 \mid x)$.

Problem 4 (AMS 207):

To obtain a robust linear regression model, we consider an error distribution given by a double exponential, instead of a normal. As for the Student's t-distribution case, it is possible to use the fact that the double exponential is a scale mixture of normals to obtain an efficient Gibbs sampler for exploring the posterior distribution of the model parameters. More specifically, consider the model

$$Y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}; \quad Y, \boldsymbol{\varepsilon} \in \mathbb{R}^n,\ X \in \mathbb{R}^{n\times p},\ \boldsymbol{\beta} \in \mathbb{R}^p$$

where

$$p(\boldsymbol{\varepsilon} \mid \sigma^2) = \prod_{i=1}^{n} \frac{1}{2\sqrt{\sigma^2}} e^{-|\varepsilon_i|/\sqrt{\sigma^2}}\ .$$

Consider a non-informative prior $p(\boldsymbol{\beta}, \sigma^2) \propto 1/\sigma^2$.

1. (20%) Use the fact that

$$\frac{1}{2\sqrt{\sigma^2}} e^{-|\varepsilon_i|/\sqrt{\sigma^2}} = \int_0^\infty \left(\frac{1}{\sqrt{2\pi}\sqrt{s^2}} e^{-\varepsilon_i^2/(2s^2)} \right) \left(\frac{1}{2\sigma^2} e^{-s^2/(2\sigma^2)} \right) ds^2$$

   to write the likelihood as a scale mixture of normals.

2. (35%) Introduce latent variables in order to write the model as a hierarchical normal linear model.

3. (45%) Obtain and identify the posterior full conditional distributions for all the parameters of the hierarchical model.

Problem 5 (AMS 256):

1. (50%) Consider the model

$$y_i = \alpha + \sin(x_i + \beta) + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} N(0, \sigma^2), \quad i = 1, ..., n. \tag{5.1}$$

Here $y_1, ..., y_n$ are the responses and $x_1, ..., x_n$ are the predictors. Assume $\sum_{i=1}^n \sin(x_i) = \sum_{i=1}^n \cos(x_i) = \sum_{i=1}^n \sin(2x_i) = 0$.

(a) (20%) Represent (5.1) in the form of a linear model with vector of regression coefficients $(\alpha, \cos(\beta), \sin(\beta))$. (*Hint:* $\sin(A + B) = \sin(A)\cos(B) + \cos(A)\sin(B)$.)

(b) (30%) Provide the explicit form of the least squares estimators of the parameters in the linear model representation.

2. (50%) Consider the two-way ANOVA model $y_{ij} = \mu + \zeta_i + \lambda_j + \epsilon_{ij}$, $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$, $i = 1, ..., 3$ and $j = 1, ..., 4$.

(a) (30%) Express the above equations in the form of $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$, and specify $\boldsymbol{y}$, $\boldsymbol{X}$ and $\boldsymbol{\gamma}$. What is the rank($\boldsymbol{X}$)? How many constraints are required to make the predictor matrix full column rank?

(b) (20%) Is $\zeta_1 - \zeta_3 + \lambda_4 - \lambda_1$ estimable?