Include **all your code and all outputs** from R. Submit your solutions on Blackboard or through email, by 11:59pm Friday July 24th . You are to work completely independently on this exam; however you may use notes, your textbook, etc.

1. (40 points) Consider finding $\sigma^2 = E(X^2)$ when X has the density that is proportional to

$$q(x) = \frac{e^x}{e^{2x}+1}, x \in (-\infty, \infty).$$

(a) Find the Monte Carlo estimate of σ^2 using acceptance-rejection sampling. Take the candidate distribution as a double exponential, i.e. $g = exp(-|x|)$.
(b) Find the normalizing constant of the pdf by integrating $q(x)$ over the support. Then derive the CDF of X.
(c) Generate a sample of X using inverse transform method and find the Monte Carlo estimate of σ^2.
(d) Repeat the estimation using importance sampling with standardized weights.

2. (10 points) Consider the inverse Gaussian distribution with density

$$f(x|\theta_1, \theta_2) \propto x^{-1.5} exp\{-\theta_1 x - \frac{\theta_2}{x} + \psi(\theta_1, \theta_2)\}, \text{ where } \psi(\theta_1, \theta_2) = 2\sqrt{\theta_1\theta_2} + log(2\theta_2)$$

Estimate the E(X) using MCMC. You may take the proposal distribution as a Gamma($\sqrt{\theta_2/\theta_1}$, 1).

3. (30 points) Consider the data on coal-mining disasters from 1851 to 1962 (coal.txt data on blackboard). The rate of accidents per year appears to decrease around 1900, so we consider a change-point model for these data. Let X_j be the number of accidents in year j. $X_j \sim Poisson(\lambda_1), j = 1, ..., \theta$, and $X_j \sim Poisson(\lambda_2), j = \theta+1, ..., 112$. The change-point occurs after the θ th year in the series. This model has parameters are θ, λ_1, λ_2. Below are three sets of priors for a Bayesian analysis of this model. Assume prior $\lambda_i \sim Gamma(3, 1)$ for $i = 1, 2$, and assume θ follows a discrete uniform distribution over $\{1, ..., 111\}$.

(a) Derive the posterior distribution of (θ, λ_1, λ_2).
(b) Derive the conditional posterior distributions necessary to carry out Gibbs sampling for this change-point model. Note the conditional posterior of θ is a discrete distribution with

$$P(\theta = k|\cdot) = e^{k(\lambda_2-\lambda_1)}(\lambda_1/\lambda_2)^{\sum_{i<k} x_i}.$$

(c) Implement the Gibbs sampler. Use a suite of convergence diagnostics to evaluate the convergence and mixing of your sampler.

4. (20 points) Compare bootstrapped CIs for the population 90th percentile to the large sample estimate as in the notes for (a) $Exp(1)$ data, (b) $N(0, 1)$ data, (c) $U(0, 1)$ data, and (d) $\chi^2(1)$ data. For sample sizes of $n = 100$ and replicate $B = 500$.

(a) Compute coverage probabilities of the two intervals and average interval length. (You need to run the intervals for M times.)
(b) Summarizing your results in a table. Comment on your findings. Which is better?