AUBURN UNIVERSITY

Stat 7650 - Computational Statistics

In-Class Component - Final Exam May 5, 2025

Duration: 120 minutes

Instructions:

1. This is the in-class component of the final exam. If you want to ask a question, please raise your hand, but no answer is guaranteed unless there is an ambiguity or typo in the question.
2. Calculators are permitted but must be supplied by you.
3. Work all the problems. Please submit only the solutions you want to be grade (i.e., do not submit any scratch work).
4. You need to start each problem on a new page. Clearly label each problem and write your name at top right of each page.
5. You must always **explain your answers** and **show your work** to receive **full credit** or proper **partial credit**. If necessary, you can use both sides of the sheets, but make sure you have indicated doing so.
6. There are 6 questions. Please make sure you have the pages with all 6 questions.

QUESTIONS

Q1. Use Laplace's method to approximate the expression in $\binom{2n}{n}$ as $n \to \infty$.
Hint: Recall the approximation we derived for the factorial using Laplace's method in class.

Q2. Consider the Weibull distributions defined below.

(a) Consider first the *standard Weibull distribution* with probability density function

$$f_0(x_0) = \beta x_0^{\beta-1} e^{-x_0^\beta}, \quad x_0 > 0,\ \beta > 0$$

and cumulative distribution function

$$F_0(x_0) = 1 - e^{-x_0^\beta}.$$

Explain how the *inversion method* can be used to generate samples from the standard Weibull distribution f_0.

(b) Now consider the *general Weibull distribution* with density function

$$f(x) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-(x/\lambda)^\beta}, \quad x > 0,\ \beta > 0,\ \lambda > 0.$$

Show that if $x_0 \sim f_0$, then the transformation $x = \lambda x_0$ has the general Weibull density $f(x)$. Discuss how this result can be used to generate random variables from the general Weibull distribution.

Q3. Consider a triangular distribution with the density function

$$f(x) = 1 - |x|, \quad x \in [-1, 1].$$

(a) Propose an accept-reject procedure to simulate a random variable X having the triangular distribution described above. (*Hint: Keep it simple in choosing your envelope density! Don't forget to write the full algorithm with steps.*)

(b) What is the acceptance probability for your proposed method?

(c) Show how to simulate directly from the triangular distribution using the inverse cumulative distribution function (cdf) method. (Again, don't forget to write the full algorithm with steps.)

Q4. Let $X_1, \dots, X_n$ be i.i.d. from a mixture density

$$P_\theta(x) = \theta f(x) + (1-\theta) g(x), \quad \theta \in [0,1],$$

where f and g are known density functions, i.e., θ is the only unknown parameter. Derive the EM algorithm to find the maximum likelihood estimator of θ. (*Hint: You need to introduce "missing data"—we've done this in class and in homework.*)

Q5. Variance Reduction via Importance Sampling

Suppose we are interested in estimating the following integral:

$$\mu = \int_0^1 \cos\left(\frac{\pi x}{2}\right) dx = \frac{2}{\pi}.$$

(a) Rewrite the integral μ as an expectation with respect to the uniform distribution on $[0,1]$, and describe how a Monte Carlo estimator $\widehat{\mu}_{\text{MC}}$ can be constructed using n i.i.d. samples from Uniform$(0,1)$.

(b) Define an importance sampling estimator $\widehat{\mu}_{\text{IS}}$ for this same integral using the proposal density

$$g(x) = \frac{3}{2}(1 - x^2), \quad x \in [0,1].$$

Show the expression for the importance weights and write the full estimator.

(c) Explain in a few sentences why this choice of $g(x)$ can lead to significant variance reduction compared to uniform sampling. What property of $g(x)$ makes it efficient in this context?

(d) Briefly define the concept of *Effective Sample Size (ESS)* in the context of importance sampling. What does a high ESS value indicate about the quality of the proposal distribution?

Q6. Gibbs Sampling and Conditional Distributions

Consider the following hierarchical normal model:

- Data: $X_i \sim \mathcal{N}(\theta_i, 1)$, independently for $i = 1, \ldots, n$,
- Prior: $\theta_i \mid \sigma^2 \sim \mathcal{N}(0, \sigma^2)$,
- Hyperprior: $\sigma^{-2} \sim \text{Gamma}(a, b)$, where this is the *rate parameterization* of the Gamma distribution.

The probability density functions (pdfs) are given by:

- Normal distribution with mean μ and variance σ^2:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

- Gamma distribution with shape $a > 0$ and *rate* $b > 0$:

$$f(x \mid a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x > 0$$

(a) Derive the full conditional distribution of $\theta_i \mid x_i, \sigma^2$, and explain why it has a normal form.

(b) Derive the full conditional distribution of $\sigma^{-2} \mid \mathbf{x}, \boldsymbol{\theta}$, and explain why it follows a Gamma distribution.

(c) Write *pseudo-code* (not actual R code) for implementing the Gibbs sampler for this model, clearly indicating the steps for updating each parameter.

(d) Suppose the goal is to estimate $\|\boldsymbol{\theta}\|^2 = \sum_{i=1}^n \theta_i^2$. Describe how you would use the Gibbs sampler output to estimate this quantity. Also describe and discuss how to implement and the benefit of Rao-Blackwellization in this context.

AUBURN UNIVERSITY

Stat 7650 - Computational Statistics

Take-Home Component - Final Exam   May 5 2025

Due 12 pm on May 7, 2025

Instructions:

1. This is the take-home exam component of the Final exam. If you want to ask a question, please email me, but no answer is guaranteed unless it is to clarify the question.
2. Work all the problems. Please submit only the solutions you want me to grade (i.e., do not submit any scratch work).
3. You need to start each problem on a new page. Clearly label each problem and write your name at top right of each page.
4. You must always **explain your answers** and **show your work** to receive **full credit** or proper **partial credit**. If necessary, you can use both sides of the sheets, but make sure you have indicated doing so.
5. There are 4 questions in this part. Please make sure you have the sheet with the 2 questions.
6. Include clear, well-documented code, output, and concise but informative written explanations.

QUESTIONS

Q1. (Bayesian Inference and Hypothesis Testing for the Weibull Model)
Let $X_1, \ldots, X_n$ be an i.i.d. sample from a Weibull(λ, κ) distribution with pdf:

$$f(x \mid \lambda, \kappa) = \frac{\kappa}{\lambda}\left(\frac{x}{\lambda}\right)^{\kappa-1} \exp\left(-\left(\frac{x}{\lambda}\right)^{\kappa}\right), \quad x > 0.$$

Assume the following prior on (λ, κ):

$$\pi(\lambda, \kappa) \propto e^{-(\lambda + c\kappa)}\kappa^{b-1}, \quad \text{for constants } b > 0 \text{ and } c > 0.$$

Note: For this problem, use simulated data of size $n = 20$ from a Weibull distribution with parameters $\lambda = 1.5$, $\kappa = 2$) using `set.seed(123)`.

(a) Derive the unnormalized posterior density $\pi(\lambda, \kappa \mid \mathbf{x})$ up to a proportionality constant, given a sample $\mathbf{x} = (x_1, \ldots, x_n)$.

(b) Implement the Metropolis-Hastings algorithm in a programming language of your choice (e.g., R or Python) to sample from the posterior of (λ, κ). You may use exponential proposal distributions as discussed in class. Report and interpret the posterior mean and a 95% credible interval for both λ and κ. Also provide a trace plot and autocorrelation plots to assess convergence.

(c) Use your MCMC samples to estimate the marginal posterior distribution of κ, and produce a histogram (with density overlay) of this distribution.

(d) Based on your posterior samples, conduct an informal Bayesian test of the hypothesis $H_0 : \kappa = 1$ (i.e., that the underlying distribution is exponential). State your conclusion and support it with a numerical and graphical summary.

(e) Reflect on the impact of the prior parameters b and c on the posterior results. Briefly discuss how your findings might change under different choices for b and c.

Q2. Let (X, Y) have a uniform distribution on the unit disc $\mathcal{D} = \{(x, y) : x^2 + y^2 \leq 1\}$, i.e., the joint density function is given by:

$$f(x, y) = \begin{cases} \frac{1}{\pi}, & \text{if } (x, y) \in \mathcal{D} \\ 0, & \text{otherwise} \end{cases}$$

(a) Write down a Gibbs sampler algorithm to simulate from the joint distribution specified above; in particular, provide the full conditional distributions.

(b) Suggest another strategy to carry out this simulation using another region that contains $\mathcal{D}$. (*Hint: Remember the "Fundamental Theorem of Simulation".*)

(c) Use an AI-based tool to visualize the uniform distribution on the unit disc. Analyze the distribution's symmetry and uniformity by plotting a large number of simulated points (X, Y). Reflect on how the visual representation helps in understanding the distribution characteristics and the challenges of simulating from $\mathcal{D}$.

Q3. Let X and Y be independent $N(0, 1)$ random variables, and suppose that the goal is to approximate $P(X/Y \leq t)$, where t is a fixed number.

(a) Describe a simple or naive Monte Carlo approach to estimate $P(X/Y \leq t)$.

(b) Propose a more sophisticated Monte Carlo approach based on Rao-Blackwellization. Discuss in what sense this estimator is better than the one described in Part (a).

(c) The method based on Rao-Blackwellization described in part(b) is effective, yet an exact formula can provide a more precise solution. Consider the random variables X and Y, both independently distributed as $N(0, 1)$.

  (i) Derive the exact formula for the density function of the ratio X/Y

  OR

  (ii) Derive the cumulative distribution function (cdf) for X/Y.

  (Note that you need to do only (i) or (ii), but not both here.)

(d) Use an AI tool to generate a large number of samples for X and Y, compute the ratio X/Y for each pair, and estimate $P(X/Y \leq t)$ using the empirical distribution of these ratios. Compare your simulation results with the theoretical results obtained in part (c). Discuss the effectiveness of the AI tool in enhancing understanding and accuracy of the simulation.

Q4. (Comparative Study of Monte Carlo and Importance Sampling Estimators) In this question, you will compare the performance of naive Monte Carlo and importance sampling estimators for estimating the integral

$$\mu = \int_0^1 \sin(\pi x/2)\, dx = 2/\pi.$$

(a) **Naive Monte Carlo:** Generate $n = 10{,}000$ i.i.d. samples from the Uniform$(0,1)$ distribution. Construct the standard Monte Carlo estimator

$$\widehat{\mu}_{\text{MC}} = \frac{1}{n}\sum_{i=1}^{n} \sin(\pi X_i/2),$$

and report the estimate, estimated standard error, and a 95% confidence interval for μ.

(b) **Importance Sampling:** Recall that we used the proposal density $g(x) = \frac{3}{2}(1 - x^2),\quad x \in [0,1]$ for estimating the integral $\int_0^1 \cos(\pi x/2)\, dx$ in class. How would you obtain a good proposal density by modifying the above $g(x)$? Using your new proposal density, generate $n = 10{,}000$ samples from this distribution (you may use rejection sampling if needed). Construct the importance sampling estimator

$$\widehat{\mu}_{\text{IS}} = \frac{1}{n}\sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)} h(X_i), \quad \text{where } f(x) = 1 \text{ and } h(x) = \sin(\pi x/2),$$

and report the estimate, estimated standard error, and a 95% confidence interval for μ.

(c) **Comparison:** Compare the performance of the two estimators in terms of: (i) Accuracy (closeness to true value $2/\pi$), (ii) Precision (standard error and confidence interval width), (iii) Stability (based on variability across runs). Comment on the variance reduction achieved.

(d) **Effective Sample Size (ESS):** Compute the effective sample size (ESS) for the importance sampling estimator:

$$N_{\text{eff}} = \frac{n}{1 + s_w^2}, \quad \text{where } s_w^2 \text{ is the sample variance of the weights } w_i = \frac{1}{g(X_i)}.$$

Interpret your result: is $g(x)$ an efficient proposal?