STA531 Midterm Exam 1

Instructions

- Write your name, NetID, and signature below.
- If you need extra space for any problem, continue on the back of the page.

Community Standard

To uphold the Duke Community Standard:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself honorably in all my endeavors; and
- I will act if the Standard is compromised.

I have adhered to the Duke Community Standard in completing this exam.

Name: ______________________________

NetID: ______________________________

Signature: ______________________________

Score

1. ____________
2. ____________
3. ____________
4. ____________

Overall: ____________________

List of common distributions

$\text{Geometric}(x|\theta) = \theta(1-\theta)^x \, \mathbb{1}(x \in \{0, 1, 2, \ldots\})$ for $0 < \theta < 1$

$\text{Bernoulli}(x|\theta) = \theta^x (1-\theta)^{1-x} \, \mathbb{1}(x \in \{0, 1\})$ for $0 < \theta < 1$

$\text{Binomial}(x|n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \, \mathbb{1}(x \in \{0, 1, \ldots, n\})$ for $0 < \theta < 1$

$\text{Poisson}(x|\theta) = \dfrac{e^{-\theta}\theta^x}{x!} \, \mathbb{1}(x \in \{0, 1, 2, \ldots\})$ for $\theta > 0$

$\text{Exp}(x|\theta) = \theta e^{-\theta x} \, \mathbb{1}(x > 0)$ for $\theta > 0$

$\text{Uniform}(x|a, b) = \dfrac{1}{b-a} \, \mathbb{1}(a < x < b)$ for $a < b$

$\text{Gamma}(x|a, b) = \dfrac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \, \mathbb{1}(x > 0)$ for $a, b > 0$

$\text{Pareto}(x|\alpha, c) = \dfrac{\alpha c^\alpha}{x^{\alpha+1}} \, \mathbb{1}(x > c)$ for $\alpha, c > 0$

$\text{Beta}(x|a, b) = \dfrac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} \, \mathbb{1}(0 < x < 1)$ for $a, b > 0$

$\mathcal{N}(x|\mu, \sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\big(-\frac{1}{2\sigma^2}(x-\mu)^2\big)$ for $\mu \in \mathbb{R}$, $\sigma^2 > 0$

$\mathcal{N}(x|\mu, C) = \dfrac{1}{(2\pi)^{d/2}|C|^{1/2}} \exp\big(-\frac{1}{2}(x-\mu)^{\mathtt{T}} C^{-1} (x-\mu)\big)$ for $\mu \in \mathbb{R}^d$, $C \in \mathbb{R}^{d \times d}$ symmetric positive definite.

Exponential family form

$$p(x|\theta) = \exp\big(\varphi(\theta)^{\mathtt{T}} t(x) - \kappa(\theta)\big) h(x)$$

List of special functions

Beta function: $B(a, b) = \displaystyle\int_0^1 t^{a-1}(1-t)^{b-1} dt$ for $a, b > 0$

Gamma function: $\Gamma(x) = \displaystyle\int_0^\infty t^{x-1} e^{-t} dt$ for $x > 0$

1. (25 points) Posterior consistency and asymptotic normality

   Suppose our model is $X_1, \ldots, X_n | \theta$ i.i.d. $\sim \text{Exp}(\theta)$, along with an improper uniform prior on θ.

   (a) (Asymptotic normality) Given observed data $x_1, \ldots, x_n$, what is the asymptotic normal approximation to the posterior $p(\theta | x_{1:n})$? (Make sure you provide explicit expressions for the mean and variance in terms of $x_1, \ldots, x_n$.)

   (b) (Posterior consistency) Suppose the true distribution is $\text{Exp}(\theta_0)$, in other words, the observed data $x_1, \ldots, x_n$ are i.i.d. from $\text{Exp}(\theta_0)$. Give an informal argument that the posterior is consistent for θ_0, using the law of large numbers and the asymptotic normal approximation you derived in part 1a.

2. (25 points) Posterior predictive checks

   Suppose we have a single observation $x \in \mathbb{R}$, which we model as $X|\theta \sim \mathcal{N}(\theta, 1)$ with an improper uniform prior on θ.

   (a) What is the posterior $p(\theta|x)$?

   (b) What is the posterior predictive distribution for a replicate, $x^{\text{rep}} \in \mathbb{R}$? (Your answer should be a specific distribution with parameters depending on x, not a general formula.)

   (c) Consider the test statistic $T(x) = x$. What is the posterior predictive p-value? (Your answer should be a specific numerical value, not a general formula.)

3. (25 points) Modeling the data collection process

   (a) What is the definition of ignorability?

   (b) Give an example of a model in which the data collection process is not ignorable.

4. (25 points) Credible intervals and frequentist coverage

Suppose we have a single observation $x \in (0, 1)$ (i.e., $0 < x < 1$), which we model as $X|\theta \sim \text{Uniform}(0, \theta)$ with prior $\theta \sim \text{Uniform}(0, 1)$.

(a) Consider the interval $C(x) = [x^{0.9}, x^{0.1}]$. What is the frequentist coverage probability of this interval when the true distribution is $\text{Uniform}(0, \theta_0)$ for a given value $\theta_0 \in (0, 1)$?

(b) Show that $C(x) = [x^{0.9}, x^{0.1}]$ is an equal-tailed 80% posterior credible interval.