STA602 Exam I Spring 2022

Instructions

- This exam is **open note and open book** (it is closed to all other resources).
- **Write your name, NetID, and signature below if you use a tablet to complete the exam; otherwise put this information with question 1.**
- **Only what is on the exam will be graded (or written work submitted as one pdf file).**
- **Show all work and back up all your results for full credit.**
- **You must label/assign pages when submitting to Gradescope to avoid losing points.**
- **The exam should be submitted via Gradescope; make sure to "assign pages" as you do on your homework so we can find your solutions!**

Community Standard

By submitting my exam to Gradescope, I agree to the following:

- 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

(For TA use only — leave this section blank.)

1. ____________/8
2. ____________/2
3. ____________/7
4. ____________/11

Overall: ____________________/28

List of common distributions

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

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

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

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

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

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

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

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

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

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

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

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

1. **[8 points]** Suppose an insurer is interested in modeling insurance claims above a deductible limit of \$1000. Adopt the sampling model for claims data

$$Y_i \mid \alpha \overset{i.i.d.}{\sim} \text{Pareto}(\alpha, c = 1000), \; i = 1, \cdots, n.$$

Assume the shape parameter α has prior $\alpha \sim \text{Gamma}(a, b = \text{rate})$.

(a) **[3 points]** What is the posterior distribution of α? Show all your work.
Hint: For $c > 0$, $c^\alpha = e^{\alpha \ln c}$.

(b) **[3 points]** One can show that the maximum likelihood estimator (MLE) of α is:

$$\hat{\alpha}_n = \frac{n}{\sum_{i=1}^n \ln\left(\frac{y_i}{1000}\right)}.$$

The MLE $\hat{\alpha}_n$ is a good estimate of α based on *data alone*. Show that the posterior mean of α can be expressed as a weighted average of its prior mean and its MLE:

$$\mathbb{E}[\alpha | Y_1 = y_1, \cdots, Y_n = y_n] = w\mathbb{E}[\alpha] + (1 - w)\hat{\alpha}_n.$$

In other words, the posterior mean of α combines prior information (with weight w) and data (with weight $1 - w$).

(c) **[2 points]** What happens to the weight w if $n = 0$ and as $n \to \infty$? Interpret these two cases for the insurance problem at hand.

2. [**2 points**] State one benefit and one limitation regarding conjugate distributions. **Please provide your answer in no more than two sentences. All correct and reasonable answers will be accepted**.

3. [**7 points**] **Normal Distributions** Assume

$$Y_i \mid \theta \overset{iid}{\sim} \text{Normal}(\theta, \sigma^2), \quad i = 1, \ldots, n. \tag{0.1}$$

For now, assume that θ is unknown and $\sigma^2 > 0$ is known.

(a) [**2 points**]

i. What is a conjugate prior for θ? State the distribution and parameters in terms of μ_o, and τ_o^2.

ii. Provide the posterior distribution of $\theta \mid y_{1:n}$. **You may state it without proof**.

(b) [**2 points**] Let's now consider predicting a new observation $\tilde{Y}$ from the population after having observed the data $y_{1:n}$ and assuming the model in equation 0.1. Recall that $\sigma^2 > 0$ is known. Find the posterior predictive distribution $p(\tilde{y} \mid y_{1:n})$. **You must provide step by step details to receive full credit.**
Hint:

$$\tilde{Y} \mid \theta, \sigma^2 \sim \text{Normal}(\theta, \sigma^2) \iff \tilde{Y} = \theta + \epsilon, \quad \text{where} \quad \tilde{\epsilon} \mid \theta, \sigma^2 \sim \text{Normal}(0, \sigma^2).$$

(c) i. [**1 point**] Provide an example, where it would make sense to fit the model in equation 0.1 (give a concrete example). **Limit yourself to two–three sentences.**

ii. [**1 point**] Now give a concrete example where it seems reasonable that we do not know θ or $\sigma^2 > 0$. **Limit yourself to two–three sentences.**

(d) [**1 point**] What model would you utilize/propose using in (c), part ii? **You may either state the name of the model or give its full parameterization.**

4. **[11 points] Understanding the aging process of laptops** Suppose you are a statistical consultant working with a warranty company. You are modeling the time-to-failure of n products (denoted as $Y_1, \cdots, Y_n$). Based upon the empirical distribution of the data, you adopt the following Weibull sampling model:

$$Y_i|\gamma \overset{i.i.d.}{\sim} \text{Weibull}\big(\gamma, k\big), \quad i = 1, \cdots, n.$$

Here, $\gamma = \frac{1}{\lambda}$ is an unknown rate parameter, and k is a shape parameter (assumed fixed).

The Weibull distribution for a **single** $y_i \mid \gamma = \frac{1}{\lambda}$ can be written as the following:

$$p(y_i \mid \gamma = \frac{1}{\lambda}, k) = k\lambda(y_i\lambda)^{k-1} \exp\{-(y_i\lambda)^k\}, \quad \text{where} \quad y \geq 0.$$

You may find that https://en.wikipedia.org/wiki/Weibull_distribution may help you on part a.

For this entire problem, you must show all of your step by step derivations in order to receive full/partial credit.

(a) [**2 points**] Suppose the product under warranty is a laptop computer. What should you set k as in the Weibull model? Explain briefly.

(b) [**1 point**] Show (step by step) the following:

$$p(y_{1:n} \mid \lambda) \propto \lambda^n \lambda^{n(k-1)} \exp\Big\{ -\sum_{i=1}^{n} (y_i\lambda)^k \Big\}.$$

(c) [**1 point**] Suppose $k = 2$. Is $\lambda \sim Gamma(a = \text{shape}, b = \text{rate})$ conjugate for the likelihood function? Specifically, consider

$$p(\lambda) \propto \lambda^{a-1} \exp\{-b\lambda\}.$$

(d) **[2 points]** Let's define a new parameter such that $\lambda = \theta^{1/k}$. Show (step by step) that $\theta \sim \text{Gamma}(a, b)$ is conjugate under this re-parametrized model for any k. That is, derive the posterior distribution

$$p(\theta | y_{1:n}).$$

(e) [**3 points**] From part (d), what is the posterior mean $\mathbb{E}(\theta|y_{1:n})$? Suppose a good "data-based" estimate of θ is $\hat{\theta} = n/\sum_{i=1}^{n} y_i^k$. Write the posterior mean of θ as a weighted average of $\hat{\theta}$ and the prior mean $\mathbb{E}(\theta)$. Give two scenarios in which one would place higher weight on the prior mean.

(f) [**2 points**] Suppose you set $k = 1$ in part (d). Let Z be the time-to-failure of a new product, following $Z|\theta \sim Weibull(\theta^{1/k}, k) = Weibull(\theta, 1)$. What is the predictive distribution $p(Z|y_{1:n})$?
Hint: Let A be a random variable with probability density function $p_A(x)$. Then $B = A - c$ is a r.v. with p.d.f. $p_B(x) = p_A(x + c)$.