Total Points: 60

PSTAT 120B / **FINAL EXAMINATION** / Summer 2024 Instructor: **Ethan P. Marzban**

Name: ______________________________ **NetID:** ____________________
(*First and Last*) (*NOT Perm Number*)

Your Section: **2pm (Hyuk-Jean)** **3pm (Hyuk-Jean)** **4pm (Minwoo)** **5pm (Minwoo)**
(*Circle One*)

Instructions:

- You will have **2 hours** to complete this exam.
  - Nobody will be permitted to leave the exam room during the last 10 minutes of the exam.
- Please remember to write your name and NetID (not perm number) **at the top of each sheet** of this exam.
- You are allowed the use of **two 8.5 $\times$ 11-inch** sheets, front and back, of notes, along with the use of a **calculator**; the use of any and all other resources (including, but not limited to laptops, cell phones, textbooks, etc.) is prohibited.
- Unless otherwise specified, all multiple choice questions have **only one** correct answer.
  - Partial credit is **not** available on multiple choice questions.
- Unless otherwise specified, simplification is not needed; however, all integrals and infinite sums (unless otherwise specified) must be evaluated.
- **Good Luck!!!**

Honor Code: In signing my name below, I certify that all work appearing on this exam is entirely my own and not copied from any external source. I further certify that I have not received any unauthorized aid while taking this exam.

$\times$ ______________________________

1 Multiple Choice Questions

Please fill in the bubble(s) **on the exam below** *corresponding to your answer. You do not need to submit any additional work for these questions. No partial credit is available for multiple choice questions.*

1. (1 point) **True or False:** Given a sample $\vec{Y} = \{Y_i\}_{i=1}^n$ from a distribution with unknown parameter θ, the entire sample $\vec{Y}$ is a sufficient statistic for θ.

   √ **True** ◯ False

   Solution: As was discussed in lecture, the entire sample is a sufficient statistic for θ. You can also verify this using the Factorization Theorem.

2. (1 point) Consider $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, 1)$, where $\mu \in \mathbb{R}$ is unknown. Is it possible to find an estimator $\widehat{\mu}_n$ for μ whose variance is smaller than the Cramér-Rao Lower Bound? (Read the problem statement carefully!)

   √ **Yes, it is possible** ◯ No, it is not possible

   Solution: Note that the problem makes no mention of unbiasedness - the CRLB only applies to *unbiased* estimators. As such, it *is* possible to find a *biased* estimator whose variance is smaller than the CRLB. (See Page 25 of the Topic 4 slides, for example.)

3. (1 point) Consider an i.i.d. sample $\vec{Y} := \{Y_i\}_{i=1}^n$ from a distribution with unknown parameter θ. Suppose that $U := \sum_{i=1}^n \sqrt{Y_i}$ is a sufficient statistic for θ. Which of the following must also be a sufficient statistic for θ? (Remember, there is only one correct answer.)

   ◯ $\sum_{i=1}^n Y_i$
   √ $2\sum_{i=1}^n \sqrt{Y_i}$
   ◯ $\frac{1}{n}\sum_{i=1}^n Y_i$
   ◯ Y_1
   ◯ None of the above

   Solution: Scaling a sufficient statistic by a constant will yield another sufficient statistic.

4. (1 point) Which of the following properties does the maximum likelihood estimator **not** possess in general?

- ◯ Asymptotic normality
- ◯ Asymptotic efficiency
- ◯ Consistency
- √ **Unbiasedness (for any sample size)**
- ◯ None of the above

Solution: By the theorem "Asymptotic MLE Result", we have that maximum likelihood estimators are asymptotically normal and asymptotically efficient. It was mentioned in lecture that maximum likelihood estimators are also consistent (I actually gave half a point back to people who answered this, since it wasn't explicitly written in the lecture slides. But, it was definitely discussed at one point, which is why I couldn't award full points for that answer choice).

We saw several examples of biased MLEs, meaning the correct answer choice is "Unbiasedness (for any sample size)" - MLEs are *not* unbiased, in general.

5. (1 point) If $(Y \mid X = x) \sim$ Exp(x) and $X \sim$ Gamma$(3, 2)$, what is the correct value of $\mathbb{E}[Y]$?

- ◯ 1.500
- √ 6.000
- ◯ x [this is a lowercase x]
- ◯ X [this is an uppercase X]
- ◯ None of the above

Solution: By properties of the Exponential distribution, $\mathbb{E}[Y \mid X = x] = x$ meaning $\mathbb{E}[Y \mid X] = X$. By peroperties of the Gamma distribution, $\mathbb{E}[X] = 3 \cdot 2 = 6$. By the Law of Iterated Expectations,

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[X] = 6$$

2 Free-Response Questions

1. (6 points) Let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, and define $U := Y_2/Y_1$. Show that $F_U(u)$, the CDF (cumulative distribution function) of U, is given by

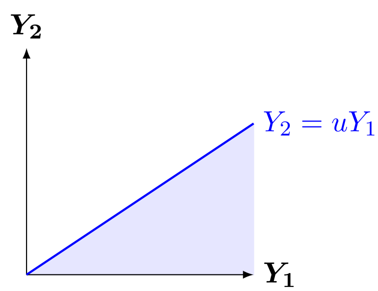
$$F_U(u) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{u}{u+1} & \text{if } u \geq 0 \end{cases}$$

For full credit, you must set up and evaluate a double integral corresponding to $F_U(u)$, and you **must sketch the region of integration**.

Solution: First note that $S_U = [0, \infty)$; hence, for any $u < 0$ we have $F_U(U) = 0$. As such, fix a $u \geq 0$. By definition, the CDF of U at u is given by

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}\left(\frac{Y_2}{Y_1} \leq u\right) = \mathbb{P}(Y_2 \leq uY_1)$$

This probability can be computed by double-integrating the joint density over the following region:



Furthermore, since $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$ we know that

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1) f_{Y_2}(y_2) = \frac{1}{\theta^2} e^{-(y_1+y_2)/\theta} \cdot \mathbb{1}_{\{y_1 \geq 0\ ,\ y_2 \geq 0\}}$$

Either order of integration is fine. Given that the joint has terms like e-to-the-negative-something, it will be easiest to include as many infinities in the bounds of integration as possible; as such, using the order dy_1 dy_2 may be slightly easier.

$$\begin{aligned} F_U(u) = \mathbb{P}(Y_2 \leq uY_1) &= \int_0^\infty \int_{y_2/u}^\infty \frac{1}{\theta^2} e^{-(y_1+y_2)/\theta} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\ &= \frac{1}{\theta} \int_0^\infty e^{-y_2/\theta} \left(\int_{y_2/u} \frac{1}{\theta} e^{-y_1/\theta} \, \mathrm{d}y_1 \right) \mathrm{d}y_2 \\ &= \frac{1}{\theta} \int_0^\infty e^{-y_2/\theta} e^{-(y_2/u)/\theta} \, \mathrm{d}y_2 = \frac{1}{\theta} \int_0^\infty e^{-\left(\frac{1+\frac{1}{u}}{\theta}\right) \cdot y_2} \, \mathrm{d}y_2 \\ &= \frac{1}{\theta} \int_0^\infty e^{-\left(\frac{u+1}{u\theta}\right) \cdot y_2} \, \mathrm{d}y_2 \end{aligned}$$

$$= \frac{1}{\cancel{\theta}} \cdot \frac{u\cancel{\theta}}{u+1} = \frac{u}{u+1}$$

Hence, putting everything together,

$$F_U(u) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{u}{u+1} & \text{if } u \geq 0 \end{cases}$$

2. Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} f(y;\theta)$ where

$$f(y;\theta) = \theta y^{\theta-1} \cdot \mathbb{1}_{\{0<y<1\}}$$

and $\theta > 0$ us an unknown parameter.

(a) (3 points) Show that $U := \left(\prod_{i=1}^{n} Y_i\right)$ is a sufficient statistic for θ.

Solution: Since our sample is stated to be i.i.d., our likelihood is just the product of the marginal densities:

$$\begin{aligned}
\mathcal{L}_{\vec{\mathbf{Y}}}(\theta) &= \prod_{i=1}^{n} f(Y_i;\theta) = \prod_{i=1}^{n} \left[\theta Y_i^{\theta-1} \cdot \mathbb{1}_{\{0 \leq Y_i \leq 1\}}\right] \\
&= \underbrace{\theta^n \left(\prod_{i=1}^{n} Y_i\right)^{\theta-1}}_{:=g(\prod_{i=1}^{n} Y_i \, , \, \theta)} \times \underbrace{\prod_{i=1}^{n} \mathbb{1}_{\{0 \leq Y_i \leq 1\}})}_{h(\vec{\mathbf{Y}})}
\end{aligned}$$

Since the likelihood factorizes into the product of two functions, one depending only on $U := \prod_{i=1}^{n} Y_i$ and θ and another depending only on $\vec{\mathbf{Y}}$, the **factorization theorem** tells us that U is a sufficient statistic for θ.

(b) (3 points) Let $W_i := -\ln(Y_i)$. Use any of the methods discussed in this course to show that $W_i \sim \text{Exp}(1/\theta)$.

Solution: Either the CDF Method or the Change of Variable formula will work - I'll demonstrate using the change of variable formula. Take $g(y) = -\ln(y)$ so that $g^{-1}(w) = e^{-w}$, and

$$\left|\frac{\mathrm{d}}{\mathrm{d}w} g^{-1}(w)\right| = |e^{-w}| = e^{-w}$$

where we have dropped the absolute value signs since e^{-w} is never negative. Hence, by the change of variable formula,

$$\begin{aligned}
f_{W_i}(w) &= f_Y(g^{-1}(w)) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}w} g^{-1}(w)\right| \\
&= \theta(e^{-w})^{\theta-1} \cdot e^{-w} \cdot \mathbb{1}_{\{0<e^{-w}<1\}} = \theta e^{-w\theta} \cdot \mathbb{1}_{\{w \geq 0\}}
\end{aligned}$$

which we recognize as the PDF of the $\text{Exp}(1/\theta)$ distribution.

(c) (3 points) Show that

$$V := \left(2\theta \sum_{i=1}^{n} [-\ln(Y_i)]\right) \sim \chi^2_{2n}$$

You may use any result from class, however you must clearly state which result/s you are using. Don't skip any steps!

Solution: Here's the way I thought about this problem:

- Note that, letting W_i be defined as in part (b) above,

$$V := 2\theta \sum_{i=1}^{n} W_i$$

- We know that the sum of i.i.d. Exp(θ) variables has distribution Gamma(n, θ), meaning

$$\sum_{i=1}^{n} W_i \sim \text{Gamma}\left(n, \frac{1}{\theta}\right)$$

- If $Y \sim$ Gamma(α, β), then $(cY) \sim$ Gamma$(\alpha, c\beta)$
- Hence,

$$V = 2\theta \cdot \sum_{i=1}^{n} W_i \sim \text{Gamma}\left(n, 2\theta\frac{1}{\theta}\right) \sim \text{Gamma}\,(n, 2)$$

- We know that the χ^2_ν distribution is equivalent to the Gamma$(\nu/2, 2)$ distribution meaning the Gamma$(n, 2)$ distribution is equivalent to the χ^2_{2n} distribution. That is, $V \sim \chi^2_{2n}$.

3. Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} f(y; \beta)$ where

$$f(y; \beta) = \frac{3\beta^3}{y^4} \cdot \mathbb{1}_{\{y \geq \beta\}}$$

and $\beta > 0$ is an unknown constant. A fact you may use without proof is that $\mathbb{E}[Y_1] = 3\beta/2$.

(a) (2 points) Derive an expression for $\widehat{\beta}_{\text{MM}}$, the method of moments estimator for β.

Solution: Since we are told that the first population moment is just $\mathbb{E}[Y_1] = 3\beta/2$, we know that the method of moments estimator must satisfy

$$\frac{3\widehat{\beta}_{\text{MM}}}{2} = \overline{Y}_n \implies \boxed{\widehat{\beta}_{\text{MM}} = \frac{2}{3}\overline{Y}_n}$$

(b) (2 points) Show that $\widehat{\beta}_{\text{MM}}$ is a consistent estimator for β - don't just cite the property that method of moments estimators are typically consistent!

Solution:

- By the **Weak Law of Large Numbers**, $\overline{Y}_n \xrightarrow{\mathbb{P}} \mathbb{E}[Y_i] = \frac{3\beta}{2}$
- Therefore, combining the WLLN with the **Continuous Mapping Theorem**,

$$\frac{2}{3}\overline{Y}_n \xrightarrow{\mathbb{P}} \frac{2}{3} \cdot \mathbb{E}[Y_1] = \frac{2}{3} \cdot \frac{3\beta}{2} = \beta$$

Hence $\widehat{\beta}_{\text{MM}} \xrightarrow{\mathbb{P}} \beta$; i.e. $\widehat{\beta}_{\text{MM}}$ is a consistent estimator for β.

(c) (3 points) Derive an expression for $\mathcal{L}_{\vec{\mathbf{Y}}}(\beta)$, the likelihood of the sample $\vec{\mathbf{Y}} = \{Y_i\}_{i=1}^n$.

Solution: Since our sample is stated to be i.i.d., our likelihood is just the product of the marginal densities:

$$\mathcal{L}_{\vec{\mathbf{Y}}}(\beta) = \prod_{i=1}^{n} f(Y_i; \beta) = \prod_{i=1}^{n} \left[\frac{3\beta^3}{Y_i^4} \cdot \mathbb{1}_{\{Y_i \geq \beta\}} \right] = 3^n \beta^{3n} \cdot \prod_{i=1}^{n} \left(\frac{1}{Y_i^4} \right) \cdot \prod_{i=1}^{n} \mathbb{1}_{\{Y_i \geq \beta\}}$$

We could leave our answer like this, but note that we can simplify the product of the indicators further. We know that $\prod_{i=1}^{n} \mathbb{1}_{\{Y_i \geq \beta\}}$ is only nonzero when all of the Y_i's are greater than β, which equivalently occurs when $Y_{(1)} \geq \beta$ where $Y_{(1)}$ denotes the sample minimum. Hence:

$$\mathcal{L}_{\vec{\mathbf{Y}}}(\beta) = 3^n \beta^{3n} \cdot \prod_{i=1}^{n} \left(\frac{1}{Y_i^4} \right) \cdot \mathbb{1}_{\{Y_{(1)} \geq \beta\}}$$

(d) (2 points) Use your answer to part (c) to find $\widehat{\beta}_{\text{MLE}}$, the method of moments estimator for β.

Solution: The likelihood is nondifferentiable in β, meaning we must maximize it by inspection. Note that the only terms involving β are $\beta^{3n} \cdot \mathbb{1}_{\{Y_{(1)} \geq \beta\}}$, meaning we only need to find the β that maximizes this quantity. Since β^{3n} is an increasing function in β, it is maximized by setting β to be as large as possible. But, the term $\mathbb{1}_{\{Y_{(1)} \geq \beta\}} = \mathbb{1}_{\{\beta \leq Y_{(1)}\}}$ restricts β to be no greater than $Y_{(1)}$. Hence, the likelihood is maximized at $Y_{(1)}$; that is,

$$\text{arg}\ \max_{\beta} \left\{ \mathcal{L}_{\vec{\mathbf{Y}}}(\beta) \right\} =: \widehat{\beta}_{\text{MLE}} = Y_{(1)}$$

4. Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} f(y;\theta)$ where

$$f(y;\theta) = \frac{3y^2}{\theta^3} \cdot \mathbb{1}_{\{0 \le y \le \theta\}}$$

and $\theta > 0$ is an unknown parameter. A fact you may use, without proof, is that the CDF (cumulative distribution function) of Y_1 is given by

$$F_{Y_1}(y) = \begin{cases} 0 & \text{if } y < 0 \\ (y/\theta)^3 & \text{if } 0 \le y < \theta \\ 1 & \text{if } y \ge \theta \end{cases}$$

Additionally, define $U := Y_{(n)}/\theta$, where $Y_{(n)} := \max_{1 \le i \le n}\{Y_i\}$ denotes the sample maximum (i.e. the n^{th} order statistic).

(a) (3 points) Show that the density of $Y_{(n)}$, the n^{th} order statistic, is given by

$$f_{Y_{(n)}}(y) = \frac{3ny^{3n-1}}{\theta^{3n}} \cdot \mathbb{1}_{\{0 \le y \le \theta\}}$$

Solution: By our formula for the density of $f_{Y_{(n)}}(y)$,

$$\begin{aligned} f_{Y_{(n)}}(y) &= n\,[F_Y(y)]^{n-1} \cdot f_Y(y) \\ &= n\left(\frac{y}{\theta}\right)^{3n-3} \cdot \frac{3y^2}{\theta^3} \cdot \mathbb{1}_{\{0 \le y \le \theta\}} = \frac{3ny^{3n-1}}{\theta^{3n}} \cdot \mathbb{1}_{\{0 \le y \le \theta\}} \end{aligned}$$

(b) (3 points) Show that U has density

$$f_U(u) = 3nu^{3n-1} \cdot \mathbb{1}_{\{0 \le u \le 1\}}$$

and use the definition of a pivotal quantity to argue that U is a pivotal quantity for θ.

Solution: Note that U is a univariate transformation of $Y_{(n)}$. Hence, we can use either the CDF method or the Change of Variable formula - I'll demonstrate using the change of variable formula. Take $g(y) = y/\theta$ so that $g^{-1}(u) = u\theta$ and

$$\left|\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u)\right| = |\theta| = \theta$$

where we have dropped the absolute values since $\theta > 0$ is given. Hence, by the Change of Variable formula:

$$\begin{aligned} f_U(u) &= f_{Y_{(n)}}(g^{-1}(u)) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u)\right| \\ &= \frac{3n(u\theta)^{3n-1}}{\theta^{3n}} \cdot \theta \cdot \mathbb{1}_{\{0 \le (u\theta) \le \theta\}} = 3nu^{3n-1} \cdot \mathbb{1}_{\{0 \le u \le 1\}} \end{aligned}$$

(c) (4 points) Find constants a and b such that $\mathbb{P}(U < a) = \alpha/2$ and $\mathbb{P}(U > b) = \alpha/2$.

Solution: Since we have the density of U, we can compute

$$\mathbb{P}(U < a) = \int_{-\infty}^{a} f_U(u)\,\mathrm{d}u = \int_0^a 3nu^{3n-1}\,\mathrm{d}u = a^{3n}$$

Therefore,

$$a^{3n} := \left(\frac{\alpha}{2}\right) \implies \boxed{a = \left(\frac{\alpha}{2}\right)^{\frac{1}{3n}}}$$

Similarly,

$$\mathbb{P}(U > b) = 1 - \mathbb{P}(U < b) = 1 - b^{3n} \overset{!}{=} \frac{\alpha}{2} \implies \boxed{b = \left(1 - \frac{\alpha}{2}\right)^{\frac{1}{3n}}}$$

(d) (4 points) Combine your answers to parts (a) and (b) to construct a $(1-\alpha) \times 100\%$ confidence interval for θ.

Solution: Following the pivotal method outlined in lecture, we start by asserting our interval is of the form

$$\{a \leq U \leq b\} = \left\{a \leq \frac{Y_{(n)}}{\theta} \leq b\right\}$$

Since we want a $(1-\alpha) \times 100\%$ coverage probability, we know that $\mathbb{P}(a \leq U \leq b) = 1 - \alpha$. By our symmetry imposition, we want $\mathbb{P}(U < a) = \alpha/2$ and $\mathbb{P}(U > b) = \alpha/2$, meaning a and b are simply the quantities we found in part (c)! Hence, the final step is to invert our interval:

$$\left\{a \leq \frac{Y_{(n)}}{\theta} \leq b\right\} = \left\{\frac{Y_{(n)}}{b} \leq \theta \leq \frac{Y_{(n)}}{a}\right\}$$

meaning, plugging in our values of a and b from part (c) above, our $(1-\alpha) \times 100\%$ confidence interval for θ is given by

$$\boxed{\left[\frac{Y_{(n)}}{\left(1 - \frac{\alpha}{2}\right)^{\frac{1}{3n}}}, \frac{Y_{(n)}}{\left(\frac{\alpha}{2}\right)^{\frac{1}{3n}}}\right]}$$

5. As she is watching the 2024 Paris Olympics, Biyonka becomes interested in performing inference on the true average length (in minutes) of a tennis match. Among an i.i.d. sample of 25 tennis matches, she observes a sample average length of 140 minutes and a sample standard deviation of 10 minutes. Suppose she wishes to test the claim that the true average length of a tennis match is 150 minutes against a two-sided alternative, using a 5% level of significance. Assume the lengths of tennis matches are normally distributed.

   (a) (1 point) Let μ denote the true average length (in minutes) of a tennis match. State the null and alternative hypotheses.

   Solution: $H_0 : \mu = 150$ vs $\mu \neq 150$.

   (b) (1 point) Should Biyonka perform a $Z-$test or a $T-$test? Justify your answer.

   Solution: This question is essentially asking us whether we know σ (the population standard deviation) or s (the sample standard deviation). From the wording of the problem, we see that the 10 minute standard deviation Biyonka obtained was the standard deviation of times *in her sample*, meaning this is the value of s, not σ. Hence, since σ is unknown, Biyonka should perform a $\boldsymbol{T}$**−test**.

   (c) (2 points) Compute the observed value of the test statistic.

   Solution: In a two-sided $T-$ test, our test statistic is

$$\left| \frac{\overline{y}_n - \mu_0}{s/\sqrt{n}} \right| = \left| \frac{140 - 150}{10/\sqrt{25}} \right| = |-5| = 5$$

(d) (2 points) Derive an expression for the $p-$value of the observed value of the test statistic. You may leave your answer in terms of the CDF of a distribution.

Solution: Since we are performing a $T-$test with $n = 25$, we use the CDF of the t_{24} distribution:

$$p = 2F_{t_{24}}(-5)$$

(Of course, recall that there are a couple of other equivalent ways of writing this quantity!)

(e) (2 points) Suppose that the $p-$value associated with Biyonka's data is 0.01 (which is **not** to say this is the right answer to part (c) above!) State the conclusions of the test, and phrase your findings in the context of the problem.

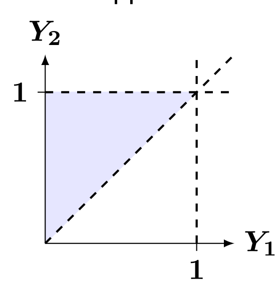
Solution: We reject the null whenever the $p-$value is smaller than the level of significance. Here, $\alpha = 0.05$ and $0.01 < 0.05$ meaning we would reject the null:

> At a 5% level of significance, there was sufficient evidence to reject the null that the true average length of a tennis match is 150 minutes, in favor of a two-sided alternative.

6. Let $(Y_1, Y_2) \sim f_{Y_1,Y_2}(y_1, y_2) = 2 \cdot \mathbb{1}_{\{0 \leq y_1 < y_2 \leq 1\}}$.

(a) (3 points) Derive an expression for $f_{Y_1}(y_1)$, the marginal density of Y_1. Be sure to include the support of Y_1 as well!

Solution: Let's start with a sketch of the support:



Firstly, this immediately tells us that $S_{Y_1} = [0, 1]$. Additionally:

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{Y_1,Y_2}(y_1, y_2) \, \mathrm{d}y_2 = \int_{\mathbb{R}} 2 \cdot \mathbb{1}_{\{0 \leq y_1 \leq 1\}} \cdot \mathbb{1}_{\{y_1 \leq y_2 \leq 1\}} \, \mathrm{d}y_2$$

$$= 2 \cdot \mathbb{1}_{\{0 \leq y_1 \leq 1\}} \cdot \int_{y_1}^{1} \mathrm{d}y_2 = 2(1 - y_1) \cdot \mathbb{1}_{\{0 \leq y_1 \leq 1\}}$$

(b) (3 points) Derive an expression for $f_{Y_1|Y_2}(y_1 \mid y_2)$, the conditional density of $(Y_1 \mid Y_2 = y_2)$. Be sure to not only include the support, but also the values of y_2 for which the density is defined! A fact you may use without proof is that the marginal density of Y_2 is given by

$$f_{Y_2}(y_2) = 2y_2 \cdot \mathbb{1}_{\{0 \leq y_2 \leq 1\}}$$

Solution: By definition, $f_{Y_1|Y_2}(y_1 \mid y_2)$ is only defined for values of y_2 such that $f_{Y_2}(y_2) \neq 0$. Since we are given the density and support of Y_2, we see that $f_{Y_2}(y_2) \neq 0$ only when $y_2 \in [0, 1]$; hence, the conditional density is defined only for $y_2 \in [0, 1]$. For such a y_2, we have

$$\begin{aligned} f_{Y_1|Y_2}(y_1 \mid y_2) &= \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} \\ &= \frac{2 \cdot \mathbb{1}_{\{0 \leq y_2 \leq 1\}} \cdot \mathbb{1}_{\{0 \leq y_1 < y_2\}}}{2y_2 \cdot \mathbb{1}_{\{0 \leq y_2 \leq 1\}}} = \frac{1}{y_2} \cdot \mathbb{1}_{\{0 \leq y_1 < y_2\}} \end{aligned}$$

(c) (3 points) Compute $\mathbb{E}[Y_1 \mid Y_2 = y_2]$.

Solution: There are a couple of ways to do this. Perhaps the easiest is to recognize, by our work from part (b), $(Y_1 \mid Y_2 = y_2) \sim \text{Unif}[0, y_2]$; hence

$$\mathbb{E}[Y_1 \mid Y_2 = y_2] = \frac{y_2}{2}$$

Of course, we could have also computed

$$\mathbb{E}[Y_1 \mid Y_2 = y_2] := \int_{\mathbb{R}} y_1 f_{Y_1|Y_2}(y_1 \mid y_2) \, \mathrm{d}y_2$$

which would have yielded the same answer.

You may use the remainder of this page for scratch work; please note that nothing written on this page will be graded.

Name:____ **NetID:**____________________

You may use the remainder of this page for scratch work; please note that nothing written on this page will be graded.

You may use the remainder of this page for scratch work; please note that nothing written on this page will be graded.