MIDTERM 1 REVIEW PROBLEMS

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

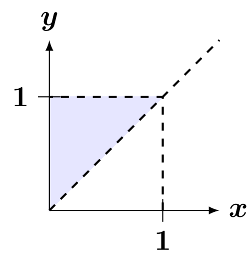
1. Let (X, Y) be a continuous bivariate random vector with joint p.d.f. given by

$$f_{X,Y}(x, y) = \begin{cases} c \cdot xy & \text{if } 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $c > 0$ is an as-of-yet undetermined constant.

(a) Find the value of c that ensures this is a valid density function.

Solution: We first sketch the support:



Either order of integration is fine:

$$\iint_{\mathbb{R}^2} f_{X,Y}(x, y) \, \mathrm{d}x = c \int_0^1 \int_0^y xy \, \mathrm{d}x \, \mathrm{d}y = c \int_0^1 \frac{1}{2} y^3 \, \mathrm{d}y = \frac{c}{8} \stackrel{!}{=} 1 \implies \boxed{c = 8}$$

(b) Find $f_Y(y)$, the marginal p.d.f. of Y.

Solution:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, \mathrm{d}x \\ &= \int_0^y 8xy \, \mathrm{d}x = 4y^3 \implies \boxed{f_Y(y) = 4y^3 \cdot \mathbb{1}_{\{y \in [0,1]\}}} \end{aligned}$$

(c) Find $f_{X|Y}(x \mid y)$, the conditional density of X given $Y = y$.

Solution:

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} \\ &= \frac{8xy \cdot \mathbb{1}_{\{0 \leq x \leq y\}} \cdot \mathbb{1}_{\{y \in [0,1]\}}}{4y^3 \cdot \mathbb{1}_{\{y \in [0,1]\}}} = \boxed{\frac{2x}{y^2} \cdot \mathbb{1}_{\{0 \leq x \leq y\}}} \end{aligned}$$

(d) Compute $\mathbb{E}[X]$ using the Law of Iterated Expectations.

Solution:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[\mathbb{E}[X \mid Y]\right] \\ \mathbb{E}[X \mid Y = y] &= \int_0^y \frac{2x^2}{y^2}\,\mathsf{d}x = \frac{2}{3}\cdot\frac{1}{y^2}\cdot y^3 = \frac{2}{3}Y \\ \mathbb{E}[X \mid Y] &= \frac{2}{3}Y \\ \mathbb{E}[Y] &= \int_0^1 4y^4\,\mathsf{d}y = \frac{4}{5} \\ \mathbb{E}[X] &= \mathbb{E}\left[\frac{2}{3}Y\right] = \frac{2}{3}\mathbb{E}[Y] = \frac{2}{3}\cdot\frac{4}{5} = \boxed{\frac{8}{15}}\end{aligned}$$

(e) Find $f_X(x)$, and verify your answer to part (d).

Solution:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathsf{d}y = \int_x^1 8xy\,\mathsf{d}y = 4x(1-x^2) \implies \boxed{f_X(x) = 4x(1-x^2)\cdot \mathbb{1}_{\{x\in[0,1]\}}}$$

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x)\,\mathsf{d}x = \int_0^1 4x^2(1-x^2)\,\mathsf{d}x = \frac{4}{3} - \frac{4}{5} = \frac{8}{15} \checkmark$$

2. In each of the following parts, you will be provided with the conditional distribution of $(X \mid Y)$ and the marginal distribution Y. Using the provided information, compute $\mathbb{E}[X]$.

(a) $(X \mid Y = y) \sim$ Bin(y, p); $\quad Y \sim$ Pois(μ)

Solution:

$$\begin{aligned}&\mathbb{E}[X \mid Y = y] = yp \\ \implies\ &\mathbb{E}[X \mid Y] = Yp \\ &\mathbb{E}[Y] = \mu \\ &\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[Yp] = p\mathbb{E}[Y] = \boxed{p\mu}\end{aligned}$$

(b) $(X \mid Y = y) \sim$ Exp(y); $\quad Y \sim$ Gamma(α, β)

Solution:

$$\begin{aligned}&\mathbb{E}[X \mid Y = y] = y \\ \implies\ &\mathbb{E}[X \mid Y] = Y \\ &\mathbb{E}[Y] = \alpha\beta\end{aligned}$$

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

3. Let $(Y_1 \mid Y_2 = y_2) \sim \text{Exp}(1/y_2)$ and $Y_2 \sim \text{Exp}(\beta)$. Find an expression for $f_{Y_1}(y_1)$, the marginal density of Y_1. Be sure to include the support of Y_1!

Solution: Since we are conditioning on a continuous random variable, we find the desired marginal density by *integrating*:

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{\mathbb{R}} f_{Y_1 \mid Y_2}(y_1 \mid y_2) \cdot f_{Y_2}(y_2) \, \mathrm{d}y_2 \\ &= \int_0^\infty y_2 e^{-y_1 y_2} \cdot \frac{1}{\beta} e^{-y_2/\beta} \, \mathrm{d}y_2 \\ &= \frac{1}{\beta} \cdot \int_0^\infty y_2 e^{-y_2/\left(\frac{\beta}{\beta y_1 + 1}\right)} \, \mathrm{d}y_2 \end{aligned}$$

Like with Problem 3 from HW01, there are two ways to solve this integral: one is to use $u-$substitutions, and the other is to multiply and divide by a constant to transform the integrand to be a proper density. I'll demonstrate the latter: we'd like to make the integrand the density of a Gamma$(2, \beta/(\beta y_1 + 1))$ distribution, as all that is missing is the normalizing constant. Therefore:

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{1}{\beta} \cdot \int_0^\infty y_2 e^{-y_2/\left(\frac{\beta}{\beta y_1 + 1}\right)} \, \mathrm{d}y_2 \\ &= \frac{1}{\beta} \cdot \Gamma(2) \cdot \left(\frac{\beta}{\beta y_1 + 1} \right)^2 \cdot \cancelto{1}{\int_0^\infty \frac{1}{\Gamma(2) \left(\frac{\beta}{\beta y_1 + 1} \right)^2} \cdot y_2^{2-1} e^{-y_2/\left(\frac{\beta}{\beta y_1 + 1}\right)} \, \mathrm{d}y_2} \\ &= \frac{\beta}{(1 + \beta y_1)^2} \end{aligned}$$

which is valid for $y_1 > 0$:

$$f_{Y_1}(y_1) = \frac{\beta}{(1 + \beta y_1)^2} \cdot \mathbb{1}_{\{y_1 \geq 0\}}$$

4. Let Y_1 and Y_2 have the joint probability density function given by

$$f_{Y_1, Y_2}(y_1, y_2) = 6(1 - y_2) \cdot \mathbb{1}_{\{0 \leq y_1 \leq y_2 \leq 1\}}$$

(a) Find the marginal density functions for Y_1 and Y_2.

Solution:

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{Y_1, Y_2}(y_1, y_2) \, \mathrm{d}y_2 - \int_{y_1}^1 6(1 - y_2) \, \mathrm{d}y_2 \implies 3(1 - y_1)^2 \cdot \mathbb{1}_{\{0 \leq y_1 \leq 1\}}$$

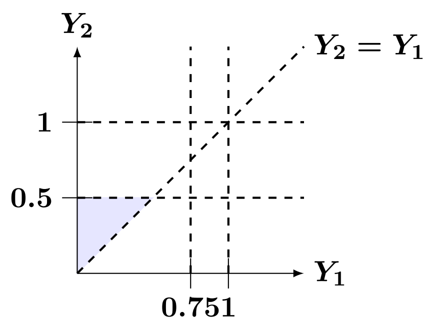
$$f_{Y_2}(y_2) = \int_{\mathbb{R}} f_{Y_1,Y_2}(y_1,y_2)\,\mathsf{d}y_1 - \int_0^{y_2} 6(1-y_2)\,\mathsf{d}y_1 \implies 6y_2(1-y_2)\cdot \mathbb{1}_{\{0\le y_2\le 1\}}$$

(b) Compute $\mathbb{P}(Y_2 \le 1/2 \mid Y_1 \le 3/4)$.

Solution: Since we are conditioning on an event with nonzero probability, we can use the definition of conditional probability:

$$\mathbb{P}(Y_2 \le 1/2 \mid Y_1 \le 3/4) = \frac{\mathbb{P}(Y_1 \le 3/4\,,\ Y_2 \le 1/2)}{\mathbb{P}(Y_1 \le 3/4)}$$

The denominator can be found by integrating the marginal for Y_1 that we found in part (a). The numerator must be computed using a double integral of the joint over the following region:



Triangular regions of integration are nice because (typically) either order of integration will be the same amount of work. As such, let's (somewhat arbitrarily) use the order $\mathsf{d}y_1\,\mathsf{d}y_2$:

$$\begin{aligned}
\mathbb{P}(Y_1 \le 3/4\,,\ Y_2 \le 1/2) &= \int_0^{1/2}\int_0^{y_2} 6(1-y_2)\,\mathsf{d}y_1\,\mathsf{d}y_2 \\
&= \int_0^{1/2} 6y_2(1-y_2)\,\mathsf{d}y_2 = 6\int_0^{1/2}(y_2 - y_2^2)\,\mathsf{d}y_2 = 6\left(\frac{1}{8} - \frac{1}{24}\right) = \frac{1}{2}
\end{aligned}$$

Additionally,

$$\mathbb{P}(Y_1 \le 3/4) = \int_0^{3/4} 3(1-y_1)^2\,\mathsf{d}y_1 = \left[(1-y_1)^3\right]_{y_1=3/4}^{y_1=0} = 1 - \frac{1}{64} = \frac{63}{64}$$

Therefore, putting everything together,

$$\mathbb{P}(Y_2 \le 1/2 \mid Y_1 \le 3/4) = \frac{\mathbb{P}(Y_1 \le 3/4\,,\ Y_2 \le 1/2)}{\mathbb{P}(Y_1 \le 3/4)} = \frac{1/2}{63/64} = \frac{32}{63}$$

(c) Find $f_{Y_1 \mid Y_2}(y_1 \mid y_2)$, and clearly specify the values of y_2 for which it is defined.

Solution: By our answer to part (a), $f_{Y_2}(y_2) = 0$ whenever $y_2 \notin [0,1]$. Hence, the conditional density $f_{Y_1\mid Y_2}(y_1 \mid y_2)$ is, by definition, defined only for $y_2 \in [0,1]$. For a fixed $y_2 \in [0,1]$, we have

$$f_{Y_1|Y_2}(y_1 \mid y_2) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_2}(y_2)}$$
$$= \frac{6(1-y_2) \cdot \mathbb{1}_{\{0 \le y_1 \le y_2\}} \cdot \mathbb{1}_{\{y_2 \in [0,1]\}}}{6y_2(1-y_2) \cdot \mathbb{1}_{\{0 \le y_2 \le 1\}}} = \boxed{\frac{1}{y_2} \cdot \mathbb{1}_{\{0 \le y_1 \le y_2\}}}$$

(d) Find $f_{Y_2|Y_1}(y_2 \mid y_1)$, and clearly specify the values of y_1 for which it is defined.

Solution: By our answer to part (a), $f_{Y_1}(y_1) = 0$ whenever $y_1 \notin [0,1]$. Hence, the conditional density $f_{Y_2|Y_1}(y_2 \mid y_1)$ is, by definition, defined only for $y_1 \in [0,1]$. For a fixed $y_1 \in [0,1]$, we have

$$f_{Y_2|Y_1}(y_2 \mid y_1) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_1}(y_1)}$$
$$= \frac{6(1-y_2) \cdot \mathbb{1}_{\{y_1 \le y_2 \le 1\}} \cdot \mathbb{1}_{\{y_1 \in [0,1]\}}}{3(1-y_1)^2 \cdot \mathbb{1}_{\{0 \le y_1 \le 1\}}} = \boxed{2 \cdot \frac{1-y_2}{(1-y_1)^2} \cdot \mathbb{1}_{\{y_1 \le y_2 \le 1\}}}$$

(e) Compute $\mathbb{P}(Y_2 \ge 3/4 \mid Y_1 = 1/2)$.

Solution: Since we are conditioning on an event with zero probability, we cannot use the definition of conditional probability. Instead, we must integrate the conditional density:

$$\mathbb{P}(Y_2 \ge 3/4 \mid Y_1 = 1/2) = \int_{3/4}^{\infty} f_{Y_2|Y_1}(y_2 \mid 1/2)\, \mathrm{d}y_2$$

Plugging in $y_1 = 1/2$ to our expression from part (d) yields

$$f_{Y_2|Y_1}(y_2 \mid y_1) = 2 \cdot \frac{1-y_2}{(1-1/2)^2} \cdot \mathbb{1}_{\{1/2 \le y_2 \le 1\}} = 8(1-y_2) \cdot \mathbb{1}_{\{1/2 \le y_2 \le 1\}}$$

and so

$$\mathbb{P}(Y_2 \ge 3/4 \mid Y_1 = 1/2) = \int_{3/4}^{\infty} f_{Y_2|Y_1}(y_2 \mid 1/2)\, \mathrm{d}y_2$$
$$= \int_{3/4}^{1} 8(1-y_2)\, \mathrm{d}y_2 = 4(1-y_2)^2 \Big]_{y_2=1}^{y_2=3/4} = \boxed{\frac{1}{4}}$$

(f) Compute $\mathbb{E}[Y_1 \mid Y_2]$.

Solution: Recall that our two-step procedure to computing $\mathbb{E}[Y_1 \mid Y_2]$ says to first compute $\mathbb{E}[Y_1 \mid Y_2 = y_2]$ and then plug in Y_2 in place of y_2. By our answer to part (c),

$$\mathbb{E}[Y_1 \mid Y_2 = y_2] := \int_{\mathbb{R}} y_1 \cdot f_{Y_1|Y_2}(y_1 \mid y_2)\, \mathrm{d}y_1 = \int_0^{y_2} y_1 \cdot \frac{1}{y_2}\, \mathrm{d}y_1 = \frac{y_2}{2}$$

Therefore, replacing y_2 with Y_2 we find

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

5. Let $X, Y \overset{\text{i.i.d.}}{\sim}$ Pois(λ).

   (a) Show that $(X + Y) \sim$ Pois(2λ).

Solution: The MGF method will be easiest:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = e^{\lambda(e^t - 1)} \cdot e^{\lambda(e^t - 1)} = e^{2\lambda(e^t - 1)}$$

which we recognize as the MGF of the Pois(2λ) distribution.

   (b) Identify the distribution of $(X \mid X + Y = n)$ for $n \in \mathbb{N}$ by name, including any/all relevant parameters

Solution: We proceed using the definition of conditional probabilities:

$$\mathbb{P}(X = x \mid X + Y = n) = \frac{\mathbb{P}(X = x \, , X + Y = n)}{\mathbb{P}(X + Y = n)}$$

The denominator can be computed using our result from part (a). To compute the numerator, note that the event $\{X = x \ , \ X + Y = n\}$ is equivalent to $\{X = x \ , \ Y = n - x\}$. Hence, repeatedly plugging into the Poisson PMF formula, we find:

$$\begin{aligned}
\mathbb{P}(X = x \mid X + Y = n) &= \frac{\mathbb{P}(X = x \, , X + Y = n)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = x, Y = n - x)}{\mathbb{P}(X + Y = n)} = \frac{\mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)}{\mathbb{P}(X + Y = n)} \\
&= \frac{e^{-\lambda} \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda} \cdot \frac{\lambda^{(n-x)}}{(n-x)!}}{e^{-2\lambda} \cdot \frac{(2\lambda)^n}{n!}} \\
&= \frac{n!}{x! \cdot (n - x)!} \cdot \frac{\lambda^x \lambda^{n-x}}{(2\lambda)^n} \\
&= \binom{n}{x} \left(\frac{1}{2} \right)^n
\end{aligned}$$

which allows us to conclude $(X \mid X + Y = n) \sim$ Bin$(n, 1/2)$.

6. The waiting time Y until delivery of a new component for an industrial operation is uniformly distributed over the interval from 1 to 5 days. The cost of this delay is given by $U = 2Y^2 + 3$. Find the probability density function for U using any of the methods we discussed in lecture.

Solution: Because the transformation $g(y) = 2y^2 + 3$ is strictly monotone over $S_Y = [1, 5]$ (which is the support of Y), we can use the Change of Variable method. (We could have also used the CDF method - I encourage you to try this on your own.) Take $g(y) = 2y^2 + 3$ so that $g^{-1}(u) = \sqrt{\frac{u-3}{2}}$, and

$$\left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| = \left| \frac{\mathrm{d}}{\mathrm{d}u} \sqrt{\frac{u-3}{2}} \right| = \frac{1}{2\sqrt{2(u-3)}}$$

and so, by the change of variable formula,

$$\begin{aligned} f_U(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| \\ &= \frac{1}{4} \cdot \mathbb{1}_{\left\{1 \leq \sqrt{\frac{u-3}{2}} \leq 5\right\}} \cdot \frac{1}{2\sqrt{2(u-3)}} = \frac{1}{8\sqrt{2(u-3)}} \cdot \mathbb{1}_{\{5 \leq u \leq 53\}} \end{aligned}$$

7. Let $Y \sim$ Unif$[0, 1]$ and define $U := aY + b$ for a constants $a > 0$ and $b \in \mathbb{R}$. Does U follow the uniform distribution? Justify your answer.

Solution: Using the CDF method, we find

$$\begin{aligned} F_U(u) &:= \mathbb{P}(U \leq u) = \mathbb{P}(aY + b \leq u) = \mathbb{P}\left(Y \leq \frac{u-b}{a}\right) = F_Y\left(\frac{u-b}{a}\right) \\ &= \begin{cases} 0 & \text{if } \frac{u-b}{a} < 0 \\ \frac{u-b}{a} & \text{if } 0 \leq \frac{u-b}{a} \leq 1 \\ 1 & \text{if } \frac{u-b}{a} \geq 1 \end{cases} \\ &= \begin{cases} 0 & \text{if } u < b \\ \frac{u-b}{a} & \text{if } b \leq u < a+b \\ 1 & \text{if } u \geq a+b \end{cases} \end{aligned}$$

Indeed, this is the CDF of another uniform distribution: specifically, $U \sim$ Unif$[b\,,\ a+b]$.

8. Suppose that Y has a gamma distribution with $\alpha = n/2$ for some positive integer $n \in \mathbb{N}$ and β equal to some specified value. Use the method of moment-generating functions to show that $W = 2Y/\beta$ has a χ^2_n distribution.

Solution:

$$\begin{aligned} M_W(t) &= M_{(2/\beta)Y}(t) = M_Y\left(\frac{2}{\beta} \cdot t\right) \\ &= \begin{cases} \left(1 - \beta \cdot \left(\frac{2}{\beta} \cdot t\right)\right)^{-n/2} & \text{if } \left(\frac{2}{\beta} \cdot t\right) < 1/\beta \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

$$= \begin{cases} (1-2t)^{-n/2} & \text{if } t < 1/2 \\ \infty & \text{otherwise} \end{cases}$$

which is inded the CDF of the χ^2_n distribution.

9. A parachutist wants to land at a target T, but she finds that she is equally likely to land at any point on a straight line (A, B), of which T is the midpoint. Find the probability density function of the distance between her landing point and the target. [**Hint:** Denote A by –1, B by $+1$, and T by 0. Then the parachutist's landing point has a coordinate X, which is uniformly distributed between -1 and $+1$. The distance between X and T is $|X|$.]

Solution: As the problem suggests, impose a coordinate system over the line (A, B) so that A is at -1, B is at $+1$, and T is at the origin. Then X, the landing coordinate, satisfies $X \sim$ Unif$[-1, 1]$, and $T := |X|$. Let's use the CDF method (since the transformation $g(x) = |x|$ is not strictly monotone over $S_X = [-1, 1]$). First note that $S_T = [0, 1]$, so fix a $t \in [0, 1]$:

$$\begin{aligned} F_T(t) &:= \mathbb{P}(T \leq t) = \mathbb{P}(|X| \leq t) \\ &= \mathbb{P}(-t \leq X \leq t) = F_X(t) - F_X(-t) \\ &= \frac{t+1}{2} - \frac{-t+1}{2} = t \end{aligned}$$

which we recognize as the pdf of the Unif$[0, 1]$ distribution: that is, $T \sim$ Unif$[0, 1]$.

10. Let $Y \sim \mathcal{N}(0, 1)$ and define the random variable U as $U := e^{\sigma Y + \mu}$ for constants $\sigma > 0$ and $\mu \in \mathbb{R}$.
   (a) Derive an expression for the density of U. **As An Aside:** the distribution of U is called the **lognormal distribution**.

   Solution: First note that $S_U = [0, \infty)$.

   We can use either the CDF method or the Change of Variable formula, since the transformation $g(y) := e^{\sigma Y + \mu}$ is strictly monotone over $S_Y = \mathbb{R}$. However, since the CDF of the normal distribution doesn't have as nice a form as other distributions, it may be slightly more direct to use the Change of Variable formula. Take $g(y) = e^{\sigma Y + \mu}$ so that

   $$g^{-1}(u) = \frac{\ln(u) - \mu}{\sigma} \implies \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| = \left| \frac{1}{\sigma u} \right| = \frac{1}{\sigma u}$$

   So, by the Change of Variable formula,

   $$f_U(u) = f_Y[g^{-1}(u)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right|$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2}\left(\frac{\ln(u)-\mu}{\sigma}\right)^2 \right\} \cdot \frac{1}{\sigma u}$$

$$\implies \boxed{\frac{1}{u\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2}\left(\frac{\ln(u)-\mu}{\sigma}\right)^2 \right\} \cdot \mathbb{1}_{\{u \geq 0\}}}$$

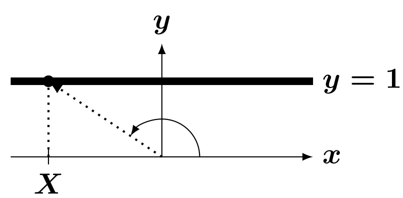
(b) We define the **median** of a continuous distribution with density $f_X(x)$ to be the value m such that $\mathbb{P}(X \leq m) = \mathbb{P}(X > m) = 1/2$. Show that the median of the lognormal distribution is e^μ. (You may use, without proof, the fact that the median of the standard normal distribution is 0.)

Solution: Essentially, what we wish to show is that $\mathbb{P}(U \leq e^\mu) = 1/2$. Directly integrating the density derived in part (a) - though valid - is *not* recommended. Instead, the trick is to rewrite things back in terms of Y:

$$\begin{aligned}
\mathbb{P}(U \leq e^\mu) &= \mathbb{P}(e^{\sigma Y + \mu} \leq e^\mu) = \mathbb{P}(\cancel{e^\mu} e^{\sigma Y} \leq \cancel{e^\mu}) \\
&= \mathbb{P}(e^{\sigma Y} \leq 1) \\
&= \mathbb{P}(\sigma Y \leq 0) = \mathbb{P}(Y \leq 0) = 1/2
\end{aligned}$$

Hence, we've shown that $\mathbb{P}(U \leq e^\mu) = 1/2$ which means, by definition, e^μ is the median of the distribution of U.

11. A particle is fired from the origin in a random (i.e. uniformly-distributed) direction pointing somewhere in the first two quadrants. The particle travels in a straight line, unobstructed, until it collides with an infinite wall located at $y = 1$. Let X denote the $x-$coordinate of the point of collision.



(a) What is the expected value of the $x-$coordinate of the point of collision? **Do NOT first find the p.d.f. of X.**

Solution: Let Θ denote the angle subtended by the trajectory of the particle, as measured from the positive $x-$axis. We can see then that

$$X = \text{cot}(\Theta)$$

Since $\Theta \sim$ Unif$[0, \pi]$ we can use the LOTUS to write

$$\begin{aligned}
\mathbb{E}[X] = \mathbb{E}\left[\text{cot}(\Theta)\right] &= \int_0^\pi \text{cot}(\theta) \cdot \frac{1}{\pi} \, \text{d}\theta \\
&= \frac{1}{\pi}\left[\int_0^{\pi/2} \text{cot}(\theta) \, \text{d}\theta + \int_{\pi/2}^{\pi} \text{cot}(\theta) \, \text{d}\theta \right]
\end{aligned}$$

Let's focus on each integral separately.

$$\int_0^{\pi/2} \text{cot}(\theta)\ \text{d}\theta = \lim_{\beta \to 0} \int_\beta^{\pi/2} \text{cot}(\theta)\ \text{d}\theta = \lim_{\beta \to 0} \text{ln}(\text{sin}\,\theta)\Big|_{\theta=\beta}^{\theta=\pi/2} = \lim_{\beta \to 0} [0 - \text{ln}(\text{sin}\,\beta)] = -\infty$$

$$\int_{\pi/2}^{\pi} \text{cot}(\theta)\ \text{d}\theta = \lim_{\beta \to \pi} \int_{\pi/2}^{\beta} \text{cot}(\theta)\ \text{d}\theta = \lim_{\beta \to \pi} \text{ln}(\text{sin}\,\theta)\Big|_{\theta=\pi/2}^{\theta=\beta} = \lim_{\beta \to \pi} [\text{ln}(\text{sin}\,\beta)] = \infty$$

Therefore, we see that $\mathbb{E}[X]$ is undefined

(b) Find $f_X(x)$, the probability density function (p.d.f.) of X

Solution: Method 1: CDF Method For an $x \in \mathbb{R}$ we have

$$F_X(x) := \mathbb{P}(X \leq x) = \mathbb{P}(\text{cot}\,\Theta \leq x) = \mathbb{P}\left(\Theta \geq \text{cot}^{-1}(x)\right) = 1 - \frac{1}{\pi}\text{cot}^{-1}(x)$$

$$f_X(x) = -\frac{\text{d}}{\text{d}x}\left(\frac{1}{\pi}\text{cot}^{-1}(x)\right) = \frac{1}{\pi(1+x^2)} \quad \text{for } x \in \mathbb{R}$$

(Note that we flipped the sign of the inequality in the first line, since $\text{cot}^{-1}(\cdot)$ is a monotonically decreasing function.)

Method 2: The Change of Variable Formula We take $g(t) = \text{cot}(t)$ so that $g^{-1}(x) = \text{cot}^{-1}(x)$ and

$$\left|\frac{\text{d}}{\text{d}x}g^{-1}(x)\right| = \left|\frac{1}{1+x^2}\right| = \frac{1}{1+x^2}$$

Since $f_\Theta(\theta) = {}^1\!/\!_\pi \cdot \mathbb{1}_{\{\theta \in [0,\pi]\}}$ we have

$$f_X(x) = \frac{1}{\pi} \cdot \mathbb{1}_{\{\text{cot}^{-1}(\theta) \in [0,\pi]\}} \cdot \frac{1}{1+x^2} = \frac{1}{\pi(1+x^2)} \cdot \mathbb{1}_{\{x \in \mathbb{R}\}}$$

As an aside: This is a special case of what is known as the **Cauchy** distribution.

(c) Confirm your answer to part (a) using your answer to part (b).

Solution: We can see that

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\ \text{d}x \text{ does not converge}$$

12. (**Challenge**) Let $X \sim \text{Exp}(1/\lambda)$, and define $Y := \lceil X \rceil$. Identify the distribution of Y **by name**, taking care to include any/all relevant parameter(s). Recall that

$$\lceil x \rceil := \text{smallest integer larger than or equal to } x$$

so, for instance, $\lceil \pi \rceil = 4$. **Hint:** Identify appropriate values for a and b such that

$$\{\lceil X \rceil = y\} = \{a < X \leq b\}$$

Also, when you are trying to identify the name of the resulting distribution, I recommend pattern-matching with the support and form of the PMF. Once you have done so, you can do a quick computation to compute $\mathbb{E}[Y]$ to help you "cheat" and factorize the PMF into a more recognizable form.

Solution: First note that the support of Y is $\{1, 2, 3, \dots\}$, meaning Y is discrete. Now, following the hint, we relate the p.m.f. of Y to the c.d.f. of X by writing

$$p_Y(y) := \mathbb{P}(Y = y) = \mathbb{P}(\lceil X \rceil = y)$$

Upon inspection, we note that

$$\{\lceil X \rceil = y\} = \{y - 1 < X \leq y\}$$

Thus, we have

$$\begin{aligned}
p_Y(y) &= \mathbb{P}(\lceil X \rceil = y) \\
&= \mathbb{P}(y - 1 < X \leq y) \\
&= F_X(y) - F_X(y - 1) \\
&= \cancel{1} - e^{-\lambda y} - \cancel{1} + e^{-\lambda(y-1)} \\
&= e^{-\lambda(y-1)} - e^{-\lambda y} \\
&= e^{-\lambda y} \cdot e^{\lambda} - e^{-\lambda y} \\
&= e^{-\lambda y}(e^{\lambda} - 1) \\
&= e^{-\lambda y} e^{-\lambda}(1 - e^{-\lambda}) \\
&= \left(e^{-\lambda}\right)^{(y-1)} (1 - e^{-\lambda}) \\
&= \left[1 - (1 - e^{-\lambda})\right]^{y-1} \cdot \left(1 - e^{-\lambda}\right)
\end{aligned}$$

showing that

$$Y \sim \text{Geom}(1 - e^{-\lambda}) \text{ on } \{1, 2, 3, \dots\}$$

As an aside: The factorization for this problem may not come very naturally to most. That is, it may be tempting to write

$$\mathbb{P}(Y = y) = e^{-\lambda y}(e^{\lambda} - 1)$$

If you have an intuition that this might follow the Geometric distribution, but don't quite know what parameter it should follow, you can "cheat" by finding the expectation of Y directly:

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{y=1}^{\infty} y \cdot e^{-\lambda y}(e^{\lambda} - 1) \\
&= (e^{\lambda} - 1) \cdot \sum_{y=1}^{\infty} y \left(e^{-\lambda}\right)^{y}
\end{aligned}$$

$$= (e^{-\lambda} - 1) \cdot \frac{e^{-\lambda}}{(1 - e^{-\lambda})^2}$$

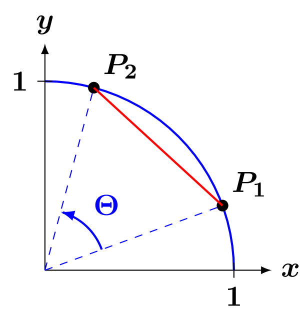
$$= \cancel{(1 - e^{-\lambda})} \cdot \frac{1}{(1 - e^{-\lambda})^{\cancel{2}}} = \frac{1}{1 - e^{-\lambda}}$$

Therefore, since the expectation of a Geometric distribution on $\{1, 2, \dots, \}$ is simply 1 divided by the parameter p, this seems to indicate that $p = 1 - e^{-\lambda}$. One can use this fact to guide the factorization of $\mathbb{P}(Y = y)$ into the more standard form of the p.m.f. of a Geometric distribution on $\{1, 2, \dots\}$.

13. (**Challenge**) Two points P_1 and P_2 are picked uniformly at random from the portion of the unit circle lying in the first quadrant. Let L denote the length of the chord connecting these two points: find $F_L(\ell)$, the density of L.

Hint: note that, since the two points are selected uniformly at random in the first quadrant, the angle Θ between radii subtended by these points is uniformly distributed between $[0, \pi/2)$. Hence, sketch a picture and use the Law of Cosines to derive a formula relating L to U; then you can use one of the methods discussed in lecture to derive the desired density.

Solution: Let's draw a picture:



The length of the red segment is L, the random variable whose density we seek. Furthermore, since the circle is a unit circle, the two blue dashed lines are both of length 1. Hence, by the Law of Cosines,

$$L^2 = 1^2 + 1^2 - 2(1)(1)\cos(\Theta) = 2(1 - \cos\Theta)$$

Because L is constrained to be positive (by virtue of being a length), we can safely take the square root of both sides to see

$$L = \sqrt{2(1 - \cos\Theta)}$$

Additionally,

$$\Theta \sim \text{Unif}[0, \pi/2)$$

meaning we are now in the realm of transformations! With a bit of work we can see that the transformation $g(\theta) = 2(1 - \cos\theta)$ is strictly monotone over $S_\Theta = [0, \pi/2)$, meaning

$$g^{-1}(\ell) = \arccos(1 - \ell^2/2)$$

Let's use the Change of Variable formula:

$$\frac{\mathrm{d}}{\mathrm{d}\ell} \text{arccos}(1 - \ell^2/2) = \frac{\ell}{\sqrt{1 - (1 - \ell^2/2)^2}}$$

meaning

$$\begin{aligned} f_L(\ell) &= f_\Theta[g^{-1}(\ell)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}\ell} g^{-1}(\ell) \right| \\ &= \frac{2}{\pi} \cdot \mathbb{1}_{\{0 \leq \text{arccos}(1 - \ell^2/2) \leq \pi/2\}} \cdot \frac{\ell}{\sqrt{1 - (1 - \ell^2/2)^2}} \\ &= \frac{2\ell}{\pi\sqrt{1 - (1 - \ell^2/2)^2}} \cdot \mathbb{1}_{\{0 \leq \ell \leq \sqrt{2}\}} \end{aligned}$$

which, if we really wanted to, can be rewritten as

$$\frac{4}{\pi\sqrt{4 - \ell^2}} \cdot \mathbb{1}_{\{0 \leq \ell \leq \sqrt{2}\}}$$