In this activity, we will explore how to compute the *probability density function* for the **product** of two continuous random variables using the **distribution method**.

To be concrete, we will consider two **independent** uniformly distributed random variables, X_1 and X_2.

- The density for X_1 we'll call $g(x)$,

$$g(x) = \begin{cases} 1 & 0 \le x \le 1, \\ \\ 0 & \text{otherwise.} \end{cases}$$

- And the density for X_2 we'll call $h(x)$,

$$h(x) = \begin{cases} \frac{1}{2} & 0 \le x \le 2, \\ \\ 0 & \text{otherwise.} \end{cases}$$

Our goal is to compute the density for the random variable X formed by multiplying X_1 and X_2,

$$X = X_1 \cdot X_2.$$

You could imagine, for example, X_1 being a random variable for current and X_2 being a random variable for resistance. According to Georg Ohm, X would then measure voltage.

1. In the space below, sketch the densities for X_1 and X_2

2. What range of values will X obtain? *Do not multiply the density values!!*

$$\boxed{\quad \le X \le \quad}$$

3. If $X = x$ and $X_1 = a$, then what value must X_2 have?

$$\boxed{X_2 = \quad\quad\quad\quad}$$

4. Given that the RVs X_1 and X_2 are **independent**, what is the value of their **joint density** (which we will temporarily call f_{X_1,X_2})? *Hint: See Equation* $4.7'$ *on page 40 of the text.*

$$\boxed{f_{X_1,X_2}(x_1, x_2) = \quad\quad\quad\quad\quad\quad}$$

5. At this point, we will take a departure from the **convolution** method used for sums of RVs. The reason for this is that the expression $\frac{x}{a}$ in (3) is not one-to-one when $x = 0$ and is undefined when $a = 0$.

   To side-step this issue, we will make use of *cumulative distribution functions.* Recall, that if the density for X is $f(x)$, then the cdf for X is

$$F(x) := P(X \le x) = \int_{-\infty}^{x} f(y)\, dy.$$

   (a) Compute the cdf for X_1, which we will call $G(x)$. Sketch the graph of $G(x)$ in the space below. Check that $G'(x) = g(x)$ everywhere that G is differentiable.

(b) Compute the cdf for X_2, which we will call $H(x)$. Sketch the graph of $H(x)$ in the space below. Check that $H'(x) = h(x)$ everywhere that H is differentiable.

6. Next, for $x \in [0, 2]$, we will express the cdf for X as follows

$$F(x) := P(X \leq x) = \int_{y=0}^{y=1} P(X_2 \leq \frac{x}{y})g(y)\, dy.$$

(a) Convince yourself that the expression above makes sense. Write a few words explaining how the integral on the right measures the probability on the left.

(b) To compute the integral, we need to carefully substitute $H(x)$ into the integrand. If you have computed $H(x)$ correctly above, you should have arrived at

$$H(x) = \begin{cases} 0 & x \geq 0, \\ \frac{x}{2} & 0 < x \leq 2, \\ 1 & \text{otherwise.} \end{cases}$$

In the integral, however, we have $H(\frac{x}{y})$. When will $H(\frac{x}{y}) = 1$? What values of y will guarantee this happens?

(c) In the previous question, you should have determined that $H(\frac{x}{y}) = 1$ whenever $y \leq \frac{x}{2}$. Keeping in mind that $g(y) = 1$ for all $y \in [0, 1]$, we arrive at

$$F(x) = \int_{y=0}^{y=1} P(X_2 \leq \frac{x}{y}) g(y)\ dy$$

$$= \int_0^1 H(\frac{x}{y})\ dy$$

$$= \int_0^{\frac{x}{2}} dy + \int_{\frac{x}{2}}^1 \frac{x}{2y}\ dy.$$

In the analysis above, we split the integral according to how $H(\frac{x}{y})$ would either equal 1 or $\frac{x}{2y}$. Implicitly, we are assuming that $x \in [0, 2]$.

(d) Evaluate the integrals above to arrive at an expression for the cdf of $X = X_1 \cdot X_2$. Since you are integrating with respect to y, you can treat x as a fixed value for the purpose of evaluating the integral. Also, remember that $\ln(1) = 0$.

If you have done everything correctly so far, you should have arrived at

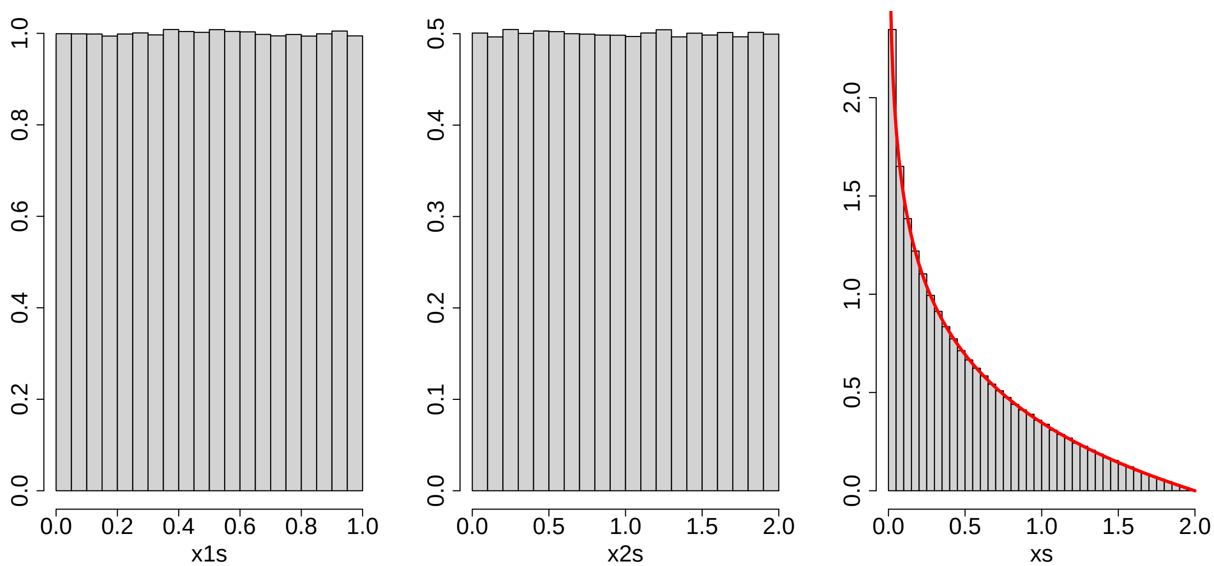
$$F(x) = \begin{cases} 0 & x \geq 0, \\ \\ \frac{x}{2}(1 - \ln(\frac{x}{2})) & 0 < x \leq 2, \\ \\ 1 & \text{otherwise.} \end{cases}$$

7. Differentiate the expressions above to arrive at a formula for the density $f(x)$ of the random variable X.

$$f(x) = \begin{cases} & x \geq 0, \\ \\ & 0 < x \leq 2, \\ \\ & \text{otherwise.} \end{cases}$$

8. Excellent! In the space below sketch the graphs of both $f(x)$ and $F(x)$.

9. I often find it helpful to "check" analytic results by running simulations in R or Python. The following histograms were created using R – sample code is available on Canvas. The red-line on the rightmost plot corresponds to the density I computed for X using the method I presented in the preceding exercises. It's a good match!



10. Does the plot on the right look like your sketch of $f(x)$? Does the solution match your intuition? Why is 0 so popular?

11. BONUS ROUND!! Some of you might be wondering why we couldn't simply log-transform the product (to turn it into a sum) and then use convolutions. Perhaps you already have a workflow in mind.

    1. $X = X_1 \cdot X_2$
    2. $\implies \log(X) = \log(X_1) + \log(X_2)$
    3. Assume independence is preserved?
    4. Convolve the densities for $\log(X_1)$ and $\log(X_2)$ to arrive at a density for $\log(X)$.
    5. Somehow recover the density for X from the density for $\log(X)$.

    Of course, we can't take the logarithm of 0, so there's that problem, bit it sounds promising for products in general. To make this plan actionable, however, we need a gadget that extract densities from transformed random variables. Here it is!

    TRANSFORMATION THEOREM – Suppose X has pdf $f_X(x)$ and form a new random variable Y by applying a monotonic function to X, that is

    $$Y = g(X)$$

    with g either strictly increasing or strictly decreasing on the set of all possible values of X (what we called D in class on Wednesday). One consequence of the monotonicity will be that g is invertible on D. Let's call this inverse function h, so that

    $$h(g(x)) = x \qquad \text{and} \qquad g(h(y)) = y.$$

    If h is differentiable (chances are good), then we can easily compute the density for the transformed variable Y.

    $$f_Y(y) = f_X(h_Y)) \cdot |h'(y)|$$

12. For practice, suppose X is uniformly distributed on $[1, 2]$. What would the density be for X^2?