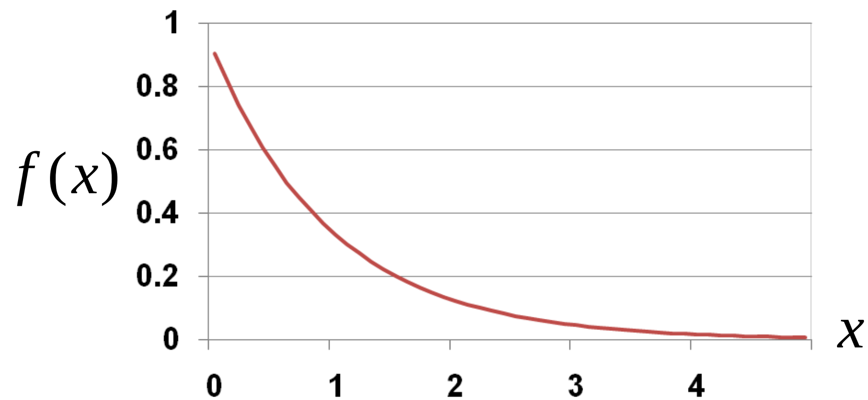
Exponential Random Variable

- X is an **Exponential RV**: $X \sim \text{Exp}(\lambda)$ Rate: $\lambda > 0$
  - Probability Density Function (PDF):

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad \text{where} \quad -\infty < x < \infty$$



  - $E[X] = \dfrac{1}{\lambda}$
  - $Var(X) = \dfrac{1}{\lambda^2}$
  - Cumulative distribution function (CDF), $F(X) = P(X \leq x)$:

$$F(x) = 1 - e^{-\lambda x} \quad \text{where } x \geq 0$$

  - Represents time until some event
    - Earthquake, request to web server, end cell phone contract, etc.

Exponential is "Memoryless"

- X = time until some event occurs
  - X ~ Exp(λ)
  - What is P(X > s + t | X > s)?

$$P(X > s + t \mid X > s) = \frac{P(X > s + t \text{ and } X > s)}{P(X > s)} = \frac{P(X > s + t)}{P(X > s)}$$

$$\frac{P(X > s + t)}{P(X > s)} = \frac{1 - F(s + t)}{1 - F(s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = 1 - F(t) = P(X > t)$$

$$\text{So, } P(X > s + t \mid X > s) = P(X > t)$$

  - After initial period of time *s*, P(X > t | •) for waiting another *t* units of time until event is same as at start
  - "Memoryless" = no impact from preceding period *s*

Visits to Web Site

- Say visitor to your web site leaves after X minutes
  - On average, visitors leave site after 5 minutes
  - Assume length of stay is Exponentially distributed
  - X ~ Exp(λ = 1/5), since E[X] = 1/λ = 5
  - What is P(X > 10)?

$$P(X > 10) = 1 - F(10) = 1 - (1 - e^{-\lambda 10}) = e^{-2} \approx 0.1353$$

  - What is P(10 < X < 20)?

$$P(10 < X < 20) = F(20) - F(10) = (1 - e^{-4}) - (1 - e^{-2}) \approx 0.1170$$

Replacing Your Laptop

- X = # hours of use until your laptop dies
  - On average, laptops die after 5000 hours of use
  - X ~ Exp(λ = 1/5000), since E[X] = 1/λ = 5000
  - You use your laptop 5 hours/day.
  - What is P(your laptop lasts 4 years)?
  - That is: P(X > (5)(365)(4) = 7300)

$$P(X > 7300) = 1 - F(7300) = 1 - (1 - e^{-7300/5000}) = e^{-1.46} \approx 0.2322$$

  - Better plan ahead... especially if you are coterming:

$$P(X > 9125) = 1 - F(9125) = e^{-1.825} \approx 0.1612 \qquad \text{(5 year plan)}$$

$$P(X > 10950) = 1 - F(10950) = e^{-2.19} \approx 0.1119 \qquad \text{(6 year plan)}$$

A Little Calculus Review

- Product rule for derivatives:

$$d(u \cdot v) = du \cdot v + u \cdot dv$$

- Derivative and integral of exponential:

$$\frac{d(e^u)}{dx} = e^u \frac{du}{dx} \qquad \int e^u du = e^u$$

- Integration by parts:

$$\int d(u \cdot v) = u \cdot v = \int v \cdot du + \int u \cdot dv$$

$$\int u \cdot dv = u \cdot v - \int v \cdot du$$

And Now, Some Calculus Practice

- Compute *n*-th moment of Exponential distribution

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$$

  - Step 1: don’t panic, think happy thoughts, recall...
  - Step 2: find *u* and *v* (and *du* and *dv*):

$$u = x^n \qquad v = -e^{-\lambda x}$$

$$du = nx^{n-1} dx \qquad dv = \lambda e^{-\lambda x} dx$$

  - Step 3: substitute (a.k.a. “plug and chug”)

$$\int u \cdot dv = \int x^n \cdot \lambda e^{-\lambda x} dx = u \cdot v - \int v \cdot du = -x^n e^{-\lambda x} + \int nx^{n-1} e^{-\lambda x} dx$$

$$E[X^n] = -x^n e^{-\lambda x} \Big|_0^\infty + \int nx^{n-1} e^{-\lambda x} dx = 0 + \frac{n}{\lambda} \int x^{n-1} \lambda e^{-\lambda x} dx = \frac{n}{\lambda} E[X^{n-1}]$$

$$\text{Base case}: E[X^0] = E[1] = 1, \quad \text{so } E[X] = \frac{1}{\lambda}, E[X^2] = \frac{2}{\lambda}\frac{1}{\lambda} = \frac{2}{\lambda^2}, \ldots$$

Discrete Joint Mass Functions

- For two discrete random variables *X* and *Y*, the **Joint Probability Mass Function** is:

$$p_{X,Y}(a,b) = P(X = a, Y = b)$$

- Marginal distributions:

$$p_X(a) = P(X = a) = \sum_y p_{X,Y}(a,y)$$

$$p_Y(b) = P(Y = b) = \sum_x p_{X,Y}(x,b)$$

- Example: X = value of die D_1, Y = value of die D_2

$$P(X = 1) = \sum_{y=1}^{6} p_{X,Y}(1,y) = \sum_{y=1}^{6} \frac{1}{36} = \frac{1}{6}$$

A Computer (or Three) in Every House

- Consider households in Silicon Valley
  - A household has C computers: C = X Macs + Y PCs
  - Assume each computer equally likely to be Mac or PC

$$P(C=c)=\begin{cases}0.16 & c=0\\0.24 & c=1\\0.28 & c=2\\0.32 & c=3\end{cases}$$

Y \ X	0	1	2	3	$p_Y(y)$
0	0.16	0.12	0.07	0.04	0.39
1	0.12	0.14	0.12	0	0.38
2	0.07	0.12	0	0	0.19
3	0.04	0	0	0	0.04
$p_X(x)$	0.39	0.38	0.19	0.04	1.00

Marginal distributions

Continuous Joint Distribution Functions

- For two continuous random variables *X* and *Y*, the **<u>Joint Cumulative Probability Distribution</u>** is:

$$F_{X,Y}(a,b) = F(a,b) = P(X \le a, Y \le b) \quad \text{where } -\infty < a,b < \infty$$

- Marginal distributions:

$$F_X(a) = P(X \le a) = \mathrm{P}(X \le a, Y < \infty) = F_{X,Y}(a,\infty)$$

$$F_Y(b) = P(Y \le b) = \mathrm{P}(X < \infty, Y \le b) = F_{X,Y}(\infty,b)$$

- Let’s look at one:

Demo

Jointly Continuous

- Random variables *X* and *Y*, are **Jointly Continuous** if there exists PDF $f_{X,Y}(x, y)$ defined over $-\infty < x, y < \infty$ such that:

$$\mathrm{P}(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x, y)\, dy\, dx$$

- Cumulative Density Function (CDF):

$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x, y)\, dy\, dx \qquad f_{X,Y}(a,b) = \frac{\partial^2}{\partial a\, \partial b} F_{X,Y}(a,b)$$

- Marginal density functions:

$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a, y)\, dy \qquad f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)\, dx$$

Multiple Integrals Without Tears

- Let *X* and *Y* be two continuous random variables
  - where $0 \le X \le 1$ and $0 \le Y \le 2$
- We want to integrate $g(x,y) = xy$ w.r.t. *X* and *Y*:
  - First, do “innermost” integral (treat *y* as a constant):

$$\int_{y=0}^{2}\int_{x=0}^{1} xy\,dx\,dy = \int_{y=0}^{2}\left(\int_{x=0}^{1} xy\,dx\right)dy = \int_{y=0}^{2} y\left[\frac{x^2}{2}\right]_0^1 dy = \int_{y=0}^{2} y\frac{1}{2}\,dy$$

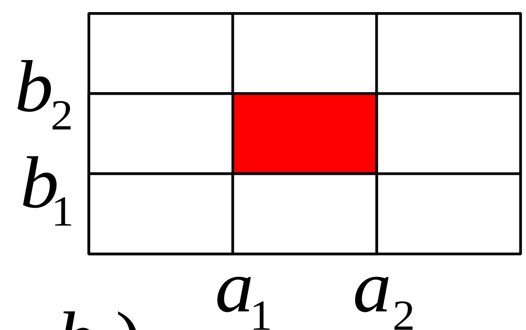
  - Then, evaluate remaining (single) integral:

$$\int_{y=0}^{2} y\frac{1}{2}\,dy = \left[\frac{y^2}{4}\right]_0^2 = 1 - 0 = 1$$

Computing Joint Probabilities

- Let $F_{X,Y}(x, y)$ be joint CDF for X and Y

$$
\begin{aligned}
\mathrm{P}(X > a, Y > b) &= 1 - P((X > a, Y > b)^c) \\
&= 1 - P((X > a)^c \cup (Y > b)^c) \\
&= 1 - P((X \le a) \cup (Y \le b)) \\
&= 1 - (P(X \le a) + P(Y \le b) - P(X \le a, Y \le b)) \\
&= 1 - F_X(a) - F_Y(b) + F_{X,Y}(a,b)
\end{aligned}
$$



$$
\begin{aligned}
&\mathrm{P}(a_1 < X \le a_2, b_1 < Y \le b_2) \\
&= F(a_2, b_2) - F(a_1, b_2) + F(a_1, b_1) - F(a_2, b_1)
\end{aligned}
$$

The Questions of Our Time

- *Y* is a <u>non-negative</u> continuous random variable
  - Probability Density Function: $f_Y(y)$
  - Already knew that:

$$E[Y] = \int_{-\infty}^{\infty} y \; f_Y(y)\, dy$$

  - But, did you know that:

$$E[Y] = \int_{0}^{\infty} P(Y > y)\, dy$$?!?

  - Analogously, in the discrete case, where $X = 1, 2, \ldots, n$

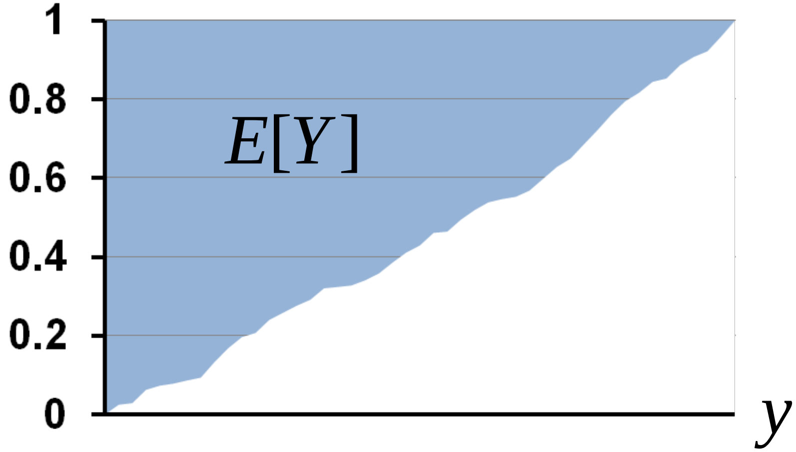
$$E[X] = \sum_{i=1}^{n} P(X \geq i)$$

Life Gives You Lemmas, Make Lemma-nade!

- A lemma in the home or office is a good thing

$$E[Y] = \int_0^\infty P(Y > y)\, dy$$

$$= \int_0^\infty (1 - F(y))\, dy$$



- Proof:

$$\int_{y=0}^\infty P(Y > y)\, dy = \int_{y=0}^\infty \int_{i=y}^\infty f_Y(i)di\, dy$$

$$= \int_{i=0}^\infty \left(\int_{y=0}^{i} dy \right) f_Y(i)di = \int_{i=0}^\infty i\ f_Y(i)di = E[Y]$$