CSE 312: Foundations of Computing II

Section 6: Continuous Random Variables Solutions

touch

1. Review of Main Concepts

(a) **Cumulative Distribution Function (cdf)**: For any random variable (discrete or continuous) X, the cumulative distribution function is defined as $F_X(x) = \mathbb{P}(X \le x)$. Notice that this function must be monotonically nondecreasing: if $x < y$ then $F_X(x) \le F_X(y)$, because $\mathbb{P}(X \le x) \le \mathbb{P}(X \le y)$. Also notice that since probabilities are between 0 and 1, that $0 \le F_X(x) \le 1$ for all x, with $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$.

(b) **Continuous Random Variable**: A continuous random variable X is one for which its cumulative distribution function $F_X(x) : \mathbb{R} \to \mathbb{R}$ is continuous everywhere. A continuous random variable has an uncountably infinite number of values.

(c) **Probability Density Function (pdf or density)**: Let X be a continuous random variable. Then the probability density function $f_X(x) : \mathbb{R} \to \mathbb{R}$ of X is defined as $f_X(x) = \frac{d}{dx} F_X(x)$. Turning this around, it means that $F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\, dt$. From this, it follows that $\mathbb{P}(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx$ and that $\int_{-\infty}^{\infty} f_X(x)dx = 1$. From the fact that $F_X(x)$ is monotonically nondecreasing it follows that $f_X(x) \ge 0$ for every real number x.

If X is a continuous random variable, note that in general $f_X(a) \ne \mathbb{P}(X = a)$, since $\mathbb{P}(X = a) = F_X(a) - F_X(a) = 0$ for all a. However, the probability that X is close to a is proportional to $f_X(a)$: for small δ, $\mathbb{P}\left(a - \frac{\delta}{2} < X < a + \frac{\delta}{2}\right) \approx \delta f_X(a)$.

(d) **i.i.d. (independent and identically distributed)**: Random variables $X_1, \ldots, X_n$ are i.i.d. (or iid) if they are independent and have the same probability mass function or probability density function.

(e) **Discrete to Continuous:**

	Discrete	Continuous
PMF/PDF	$p_X(x) = \mathbb{P}(X = x)$	$f_X(x) \ne \mathbb{P}(X = x) = 0$
CDF	$F_X(x) = \sum_{t \le x} p_X(t)$	$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
Expectation	$\mathbb{E}[X] = \sum_x x p_X(x)$	$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$
LOTUS	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

(f) **Standardizing**: Let X be any random variable (discrete or continuous, not necessarily normal), with $\mathbb{E}[X] = \mu$ and $Var(X) = \sigma^2$. If we let $Y = \frac{X - \mu}{\sigma}$, then $\mathbb{E}[Y] = 0$ and $Var(Y) = 1$.

(g) **Closure of the Normal Distribution**: Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. That is, linear transformations of normal random variables are still normal.

(h) **"Reproductive" Property of Normals**: Let $X_1, \ldots, X_n$ be independent normal random variables with $\mathbb{E}[X_i] = \mu_i$ and $Var(X_i) = \sigma_i^2$. Let $a_1, \ldots, a_n \in \mathbb{R}$ and $b \in \mathbb{R}$. Then,

$$X = \sum_{i=1}^{n} (a_i X_i + b) \sim \mathcal{N}\left(\sum_{i=1}^{n} (a_i \mu_i + b), \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

There's nothing special about the parameters – the important result here is that the resulting random variable is still normally distributed.

(i) **Law of Total Probability (Continuous):** A is an event, and X is a continuous random variable with density function $f_X(x)$.

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|X = x) f_X(x) dx$$

(j) **Central Limit Theorem (CLT)**: Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $(X_i) = \sigma^2$. Let $X = \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[X] = n\mu$ and $Var(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, which has $\mathbb{E}[\overline{X}] = \mu$ and $Var(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\overline{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0, 1)$. Similarly, as $n \to \infty$, X approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X - n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0, 1)$.

It is no surprise that $\overline{X}$ has mean μ and variance σ^2/n – this can be done with simple calculations. The importance of the CLT is that, for large n, regardless of what distribution X_i comes from, $\overline{X}$ is *approximately normally distributed with mean μ and variance σ^2/n.* Don't forget the continuity correction, only when $X_1, \ldots, X_n$ are discrete random variables.

2. Zoo of Continuous Random Variables

(a) **Uniform**: $X \sim$ Uniform(a, b) iff X has the following probability density function:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)^2}{12}$. This represents each real number from $[a, b]$ to be equally likely.

(b) **Exponential**: $X \sim$ Exponential(λ) iff X has the following probability density function:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$. $F_X(x) = 1 - e^{-\lambda x}$ for $x \geq 0$. The exponential random variable is the continuous analog of the geometric random variable: it represents the waiting time to the next event, where $\lambda > 0$ is the average number of events per unit time. Note that the exponential measures how much time passes until the next event (any real number, continuous), whereas the Poisson measures how many events occur in a unit of time (nonnegative integer, discrete). The exponential random variable X is memoryless:

$$\text{for any } s, t \geq 0,\ \mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

The geometric random variable also has this property.

(c) **Normal (Gaussian, "bell curve")**: $X \sim \mathcal{N}(\mu, \sigma^2)$ iff X has the following probability density function:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad x \in \mathbb{R}$$

$\mathbb{E}[X] = \mu$ and $Var(X) = \sigma^2$. The "standard normal" random variable is typically denoted Z and has mean 0 and variance 1: if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. The CDF has no closed form, but we denote the CDF of the standard normal as $\Phi(z) = F_Z(z) = \mathbb{P}(Z \leq z)$. Note from symmetry of the probability density function about $z = 0$ that: $\Phi(-z) = 1 - \Phi(z)$.

3. Will the battery last?

Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with expectation 10,000 miles. If the owner wants to take a 5000 mile road trip, what is the probability that she will be able to complete the trip without replacing the battery, given that the car has already been used for 2000 miles?

Solution:

Let N be a r.v. denoting the number of miles until the battery wears out. Then $N \sim \text{exp}(10,000^{-1})$, because N measures the "time" (in this case miles) before an occurrence (the battery wears out) with expectation 10,000. Since this is an exponential distribution, and the expectation of an exponential distribution is $\frac{1}{\lambda}$, $\lambda = \frac{1}{10{,}000}$. Therefore, via the property of memorylessness of the exponential distribution:

$$\mathbb{P}(N \geq 5000 | N \geq 2000) = \mathbb{P}(N \geq 3000) = 1 - \mathbb{P}(N \leq 3000) = 1 - \left(1 - e^{-\frac{3000}{10000}}\right) \approx 0.741$$

4. Create the distribution

Suppose X is a continuous random variable that is uniform on $[0, 1]$ and uniform on $[1, 2]$, but

$$\mathbb{P}(1 \leq X \leq 2) = 2 \cdot \mathbb{P}(0 \leq X < 1).$$

Outside of $[0, 2]$ the density is 0. What is the PDF and CDF of X?

Solution:

The fact that X is uniform on each of the intervals means that its PDF is constant on each. So,

$$f_X(x) = \begin{cases} c & 0 < x \leq 1 \\ d & 1 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Note that $F_X(1) - F_X(0) = c$ and $F_X(2) - F_X(1) = d$. The area under the PDF must be 1, so

$$1 = F_X(2) - F_X(0) = F_X(2) - F_X(1) + F_X(1) - F_X(0) = d + c$$

Additionally,

$$d = F_X(2) - F_X(1) = \mathbb{P}(1 \leq X \leq 2) = 2 \cdot \mathbb{P}(0 \leq X \leq 1) = 2 \cdot (F_X(1) - F_X(0)) = 2c$$

To solve for c and d in our PDF, we need only solve the system of two equations from above: $d + c = 1$, $d = 2c$. So, $d = \frac{2}{3}$ and $c = \frac{1}{3}$. Taking the integral of the PDF yields the CDF, which looks like

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{3}x & 0 < x \leq 1 \\ \frac{2}{3}x - \frac{1}{3} & 1 < x \leq 2 \\ 1 & x > 2 \end{cases}$$

5. Max of uniforms

Let $U_1, U_2, \ldots, U_n$ be mutually independent Uniform random variables on $(0, 1)$. Find the CDF and pmf for the random variable $Z = \max(U_1, \ldots, U_n)$.

Solution:

The key idea for solving this question is realizing that the max of n numbers $\max(a_1, ..., a_n)$ is less than some constant c, if and only if each individual number is less than that constant c (i.e. $a_i < c$ for all i). Using this idea, we get

$$
\begin{aligned}
F_Z(x) = \mathbb{P}(Z \le x) &= \mathbb{P}(\max(U_1, ..., U_n) \le x) \\
&= \mathbb{P}(U_1 \le x, ..., U_n \le x) \\
&= \mathbb{P}(U_1 \le x) \cdot ... \cdot \mathbb{P}(U_n \le x) && \text{[independence]} \\
&= F_{U_1}(x) \cdot ... \cdot F_{U_n}(x) \\
&= F_U(x)^n && \text{[where } U \sim \text{Unif}(0,1)\text{]}
\end{aligned}
$$

So the CDF of Z is

$$
F_Z(x) = \begin{cases} 0 & x < 0 \\ x^n & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}
$$

To find the PDF, we take the derivative of each part of the CDF, which gives us the following

$$
f_Z(x) = \begin{cases} n\, x^{n-1} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}
$$

6. Grading on a curve

In some classes (not CSE classes) an examination is regarded as being good (in the sense of determining a valid spread for those taking it) if the test scores of those taking it are well approximated by a normal density function. The instructor often uses the test scores to estimate the normal parameters μ and σ^2 and then assigns a letter grade of A to those whose test score is greater than $\mu + \sigma$, B to those whose score is between μ and $\mu + \sigma$, C to those whose score is between $\mu - \sigma$ and μ, D to those whose score is between $\mu - 2\sigma$ and $\mu - \sigma$ and F to those getting a score below $\mu - 2\sigma$. If the instructor does this and a student's grade on the test really is normally distributed with mean μ and variance σ^2, what is the probability that student will get each of the possible grades A,B,C,D and F?

Solution:

We can solve for each of these probabilities by standardizing the normal curve and then looking up each bound in the Z-table. Let X be the students score on the test. Then we have

$$
\mathbb{P}(A) = \mathbb{P}(X \ge \mu + \sigma) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \ge 1\right) = 1 - \mathbb{P}\left(\frac{X - \mu}{\sigma} < 1\right)
$$

By the closure properties of the normal random variable, $\frac{X-\mu}{\sigma}$ is distributed as a normal random variable with mean 0 and variance 1. Since this is the standard normal, we can plug it into our Φ-table to get the following:

$$
\mathbb{P}(A) = 1 - \Phi(1) = 1 - 0.84134 = 0.15866
$$

The other probabilities can be found using a similar approach:

$$
\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}(\mu < X < \mu + \sigma) = \Phi(1) - \Phi(0) = 0.34134 \\
\mathbb{P}(C) &= \mathbb{P}(\mu - \sigma < X < \mu) = \Phi(0) - \Phi(-1) = 0.34134 \\
\mathbb{P}(D) &= \mathbb{P}(\mu - 2\sigma < X < \mu - \sigma) = \Phi(-1) - \Phi(-2) = 0.13591 \\
\mathbb{P}(F) &= \mathbb{P}(X < \mu - 2\sigma) = \Phi(-2) = 0.02275
\end{aligned}
$$

7. New PDF?

Alex came up with a function that he thinks could represent a probability density function. He defined the potential pdf for X as $f(x) = \frac{1}{1+x^2}$ defined on $[0, \infty)$. Is this a valid pdf? If not, find a constant c such that the pdf $f_X(x) = \frac{c}{1+x^2}$ is valid. Then find $\mathbb{E}[X]$. (Hints: $\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$, $\tan \frac{\pi}{2} = \infty$, and $\tan 0 = 0$.)

Solution:

The area under the PDF is 1. So,

$$\int_0^\infty \frac{c}{1+x^2} dx = c \tan^{-1} x \mid_0^\infty = c\left(\frac{\pi}{2} - 0\right) = 1$$

Solving for c gives us $c = 2/\pi$. Using our value we found for c, and the definition of expectation we can compute $E[X]$ as follows:

$$\mathbb{E}[X] = \int_0^\infty \frac{cx}{1+x^2} dx = \frac{2}{\pi} \int_0^\infty \frac{x}{1+x^2} dx = \frac{1}{\pi} \ln(1+x^2) \mid_0^\infty = \infty$$

8. Throwing a dart

Consider the closed unit circle of radius r, i.e., $S = \{(x, y) : x^2 + y^2 \leq r^2\}$. Suppose we throw a dart onto this circle and are guaranteed to hit it, but the dart is equally likely to land anywhere in S. Concretely this means that the probability that the dart lands in any particular area of size A (that is entirely inside the circle of radius R), is equal to $\frac{\text{A}}{\text{Area of whole circle}}$. The density outside the circle of radius r is 0.

Let X be the distance the dart lands from the center. What is the CDF and pdf of X? What is $\mathbb{E}[X]$ and $Var(X)$?

Solution:

Since $F_X(x)$ is the probability that the dart lands inside the circle of radius x, that probability is the area of a circle of radius x divided by the area of the circle of radius r (i.e., $\pi x^2/\pi r^2$). Thus, our CDF looks like

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{r^2} & 0 < x \leq r \\ 1 & x > r \end{cases}$$

To find the PDF we just need to take the derivative of the CDF, which give us the following:

$$f_X(x) = \begin{cases} \frac{2x}{r^2} & 0 < x \leq r \\ 0 & \text{otherwise} \end{cases}$$

Using the definition of expectation we get

$$\mathbb{E}[X] = \int_{-\infty}^\infty x f_X(x) dx = \int_0^r x \frac{2x}{r^2} dx = \frac{2}{3r^2}\left(x^3 \Big|_0^r\right) = \frac{2}{3} r$$

We know that $Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^\infty x^2 f_X(x) dx = \int_0^r x^2 \frac{2x}{r^2} dx = \frac{2}{4r^2}\left(x^4 \Big|_0^r\right) = \frac{1}{2} r^2$$

Plugging this into our variance equation gives

$$Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{1}{2} r^2 - \left(\frac{2}{3} r\right)^2 = \frac{1}{18} r^2$$

9. A square dartboard?

You throw a dart at an $s \times s$ square dartboard. The goal of this game is to get the dart to land as close to the lower left corner of the dartboard as possible. However, your aim is such that the dart is equally likely to land at any point on the dartboard. Let random variable X be the length of the side of the smallest *square* B in the lower left corner of the dartboard that contains the point where the dart lands. That is, the lower left corner of B must be the same point as the lower left corner of the dartboard, and the dart lands somewhere along the upper or right edge of B. For X, find the CDF, PDF, $\mathbb{E}[X]$, and $Var(X)$.

Solution:

Since $F_X(x)$ is the probability that the dart lands inside the square of side length x, that probability is the area of a square of length x divided by the area of the square of length radius s (i.e., x^2/r^2). Thus, our CDF looks like

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ x^2/s^2, & \text{if } 0 \le x \le s \\ 1, & \text{if } x > s \end{cases}$$

To find the PDF, we just need to take the derivative of the CDF, which gives us the following:

$$f_X(x) = \frac{d}{dx}F_X(x) = \begin{cases} 2x/s^2, & \text{if } 0 \le x \le s \\ 0, & \text{otherwise} \end{cases}$$

Using the definition of expectation and variance we can compute $\mathbb{E}[X]$ and $Var(X)$ in the following manner:

$$\mathbb{E}[X] = \int_0^s x f_X(x)dx = \int_0^s \frac{2x^2}{s^2}dx = \frac{2}{s^2}\int_0^s x^2 dx = \frac{2}{3s^2}\left[x^3\right]_0^s = \frac{2}{3}s$$

$$\mathbb{E}\left[X^2\right] = \int_0^s x^2 f_X(x)dx = \int_0^s \frac{2x^3}{s^2}dx = \frac{2}{s^2}\int_0^s x^3 dx = \frac{1}{2s^2}\left[x^4\right]_0^s = \frac{1}{2}s^2$$

$$\text{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 = \frac{1}{2}s^2 - \left(\frac{2}{3}s\right)^2 = \frac{1}{18}s^2$$

10. Normal questions

(a) Let X be a normal random with parameters $\mu = 10$ and $\sigma^2 = 36$. Compute $\mathbb{P}(4 < X < 16)$.

Solution:

Let $\frac{X-10}{6} = Z$. By the scale and shift properties of normal random variables $Z \sim \mathcal{N}(0, 1)$.

$$\mathbb{P}(4 < X < 16) = \mathbb{P}\left(\frac{4-10}{6} < \frac{X-10}{6} < \frac{16-10}{6}\right) = \mathbb{P}(-1 < Z < 1) = \Phi(1) - \Phi(-1) = 0.68268$$

(b) Let X be a normal random variable with mean 5. If $\mathbb{P}(X > 9) = 0.2$, approximately what is $Var(X)$?

Solution:

Let $\sigma^2 = Var(X)$. Then,

$$\mathbb{P}(X > 9) = \mathbb{P}\left(\frac{X-5}{\sigma} > \frac{9-5}{\sigma}\right) = 1 - \Phi\left(\frac{4}{\sigma}\right) = 0.2$$

So, $\Phi\left(\frac{4}{\sigma}\right) = 0.8$. Looking up the phi values in reverse lets us undo the Φ function, and gives us $\frac{4}{\sigma} = 0.845$. Solving for σ we get $\sigma \approx 4.73$, which means that the variance is about 22.4.

(c) Let X be a normal random variable with mean 12 and variance 4. Find the value of c such that

$$\mathbb{P}(X > c) = 0.10.$$

Solution:

$$\mathbb{P}(X > c) = \mathbb{P}\left(\frac{X-12}{2} > \frac{c-12}{2}\right) = 1 - \Phi\left(\frac{c-12}{2}\right) = 0.1$$

So, $\Phi\left(\frac{c-12}{2}\right) = 0.9$. Looking up the phi values in reverse lets us undo the Φ function, and gives us $\frac{c-12}{2} = 1.29$. Solving for c we get $c \approx 14.58$.

11. Bad Computer

Each day, the probability your computer crashes is 10%, independent of every other day. Suppose we want to evaluate the computer's performance over the next 100 days.

(a) Let X be the number of crash-free days in the next 100 days. What distribution does X have? Identify $\mathbb{E}[X]$ and $Var(X)$ as well. Write an exact (possibly unsimplified) expression for $\mathbb{P}(X \geq 87)$.

Solution:

Since X counts the number of crash-free days (successes) in 100 days (trials), where each trial is a success with probability 0.9, we can see that X is binomial with $n = 100$ and $p = 0.9$, or $X \sim \text{Binomial}(100, 0.9)$. Hence, $\mathbb{E}[X] = np = 90$ and $Var(X) = np(1-p) = 9$. Finally,

$$\mathbb{P}(X \geq 87) = \sum_{k=87}^{100} \binom{100}{k} (0.9)^k (1-0.9)^{100-k}$$

(b) Approximate the probability of at least 87 crash-free days out of the next 100 days using the Central Limit Theorem. Use continuity correction.

Important: continuity correction says that if we are using the normal distribution to approximate

$$\mathbb{P}(a \leq \sum_{i=1}^{n} X_i \leq b)$$

where $a \leq b$ are integers and the X_i's are i.i.d. **discrete** random variables, then, as our approximation, we should use

$$\mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$$

where Y is the appropriate normal distribution that $\sum_{i=1}^{n} X_i$ converges to by the Central Limit Theorem.[1]

For more details see pages 209-210 in the book.

Solution:

From the previous part, we know that $\mathbb{E}[X] = 90$ and $(X) = 9$.

$$\begin{aligned}\mathbb{P}(X \geq 87) &= \mathbb{P}(86.5 < X < 100.5) = \mathbb{P}(\frac{86.5-90}{3} < \frac{X-90}{3} < \frac{100.5-90}{3}) \\ &\approx \mathbb{P}(-1.17 < \frac{X-90}{3} < 3.5) \approx \Phi(3.5) + \Phi(1.17) - 1 \approx 0.9998 + 0.8790 - 1 = 0.8788\end{aligned}$$

Notice that, if you had used $86.5 < X$ in place of $86.5 < X < 100.5$, your answer would have been nearly the same, because $\Phi(3.5)$ is so close to 1.

[1] The intuition here is that, to avoid a mismatch between discrete distributions (whose range is a set of integers) and continuous distributions, we get a better approximation by imagining that a discrete random variable, say W, is a continuous distribution with density function

$$f_W(x) := p_W(i) \quad \text{when } i - 0.5 \leq x < i + 0.5 \text{ and } i \text{ integer}$$

12. Uniform2

Alex decided he wanted to create a "new" type of distribution that will be famous, but he needs some help. He knows he wants it to be continuous and have uniform density, but he needs help working out some of the details. We'll denote a random variable X having the "Uniform-2" distribution as $X \sim$ Uniform2(a, b, c, d), where $a < b < c < d$. We want the density to be non-zero in $[a, b]$ and $[c, d]$, and zero everywhere else. Anywhere the density is non-zero, it must be equal to the same constant.

(a) Find the probability density function, $f_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$. (Hint: use a piecewise definition).

Solution:

$$f_X(x) = \begin{cases} \frac{1}{(b-a)+(d-c)}, & x \in [a, b] \cup [c, d] \\ 0, & \text{otherwise} \end{cases}$$

(b) Find the cumulative distribution function, $F_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$. (Hint: use a piecewise definition).

Solution:

$$F_X(x) = \begin{cases} 0, & x \in (-\infty, a) \\ \frac{(x-a)}{(b-a)+(d-c)}, & x \in [a, b) \\ \frac{(b-a)}{(b-a)+(d-c)}, & x \in [b, c) \\ \frac{(b-a)+(x-c)}{(b-a)+(d-c)}, & x \in [c, d) \\ 1, & x \in [d, \infty) \end{cases}$$

13. Batteries and exponential distributions

Let X_1, X_2 be independent exponential random variables, where X_i has parameter λ_i, for $1 \le i \le 2$. Let $Y = \min(X_1, X_2)$.

(a) Show that Y is an exponential random variable with parameter $\lambda = \lambda_1 + \lambda_2$. Hint: Start by computing $\mathbb{P}(Y > y)$. Two random variables with the same CDF have the same pdf. Why?

Solution:

We start with computing $\mathbb{P}(Y > y)$, by substituting in the definition of Y.

$$\mathbb{P}(Y > y) = \mathbb{P}(\min\{X_1, X_2\} > y)$$

The probability that the minimum of two values is above a value is the chance that both of them are above that value. From there, we can separate them further because X_1 and X_2 are independent.

$$\mathbb{P}(X_1 > y \cap X_2 > y) = \mathbb{P}(X_1 > y)\mathbb{P}(X_2 > y) = e^{-\lambda_1 y}e^{-\lambda_2 y}$$
$$= e^{-(\lambda_1+\lambda_2)y} = e^{-\lambda y}$$

So $F_Y(y) = 1 - \mathbb{P}(Y > y) = 1 - e^{-\lambda y}$ and $f_Y(y) = \lambda e^{-\lambda y}$ so $Y \sim$ Exp(λ), since this is the same CDF and PDF as an exponential distribution with parameter $\lambda = \lambda_1 + \lambda_2$.

(b) What is $Pr(X_1 < X_2)$? (Use the law of total probability.) The law of total probability hasn't been covered in class yet, but will be soon at which point it would be good to revisit this problem!

Solution:

By the law of total probability,

$$\mathbb{P}(X_1 < X_2) = \int_0^\infty \mathbb{P}(X_1 < X_2 | X_1 = x) f_{X_1}(x)dx = \int_0^\infty \mathbb{P}(X_2 > x)\lambda_1 e^{-\lambda_1 x}dx =$$

$$\int_0^\infty e^{-\lambda_2 x}\lambda_1 e^{-\lambda_1 x}dx = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

(c) You have a digital camera that requires two batteries to operate. You purchase n batteries, labelled $1, 2, \ldots, n$, each of which has a lifetime that is exponentially distributed with parameter λ, independently of all other batteries. Initially, you install batteries 1 and 2. Each time a battery fails, you replace it with the lowest-numbered unused battery. At the end of this process, you will be left with just one working battery. What is the expected total time until the end of the process? Justify your answer.

Solution:

Let T be the time until the end of the process. We are trying to find $\mathbb{E}[T]$. $T = Y_1 + ... + Y_{n-1}$ where Y_i is the time until we have to replace a battery from the ith pair. The reason it there are only $n-1$ RVs in the sum is because there are $n-1$ times where we have two batteries and wait for one to fail. By part (a), the time for one to fail is the min of exponentials, so $Y_i \sim Exp(2\lambda)$. Hence the expected time for the first battery to fail is $\frac{1}{2\lambda}$. By linearity and memorylessness, $\mathbb{E}[T] = \sum_{i=1}^{n-1} E[Y_1] = \frac{n-1}{2\lambda}$.

(d) In the scenario of the previous part, what is the probability that battery i is the last remaining battery as a function of i? (You might want to use the memoryless property of the exponential distribution that has been discussed.)

Solution:

If there are two batteries i, j in the flashlight, by part (b), the probability each outlasts each other is $1/2$. Hence, the last battery n has probability $1/2$ of being the last one remaining. The second to last battery $n-1$ has to beat out the previous battery and the n^{th}, so the probability it lasts the longest is $(1/2)^2 = 1/4$. Work down inductively to get that the probability the i^{th} is the last remaining is $(1/2)^{n-i+1}$ for $i \geq 3$. Finally the first two batteries share the remaining probability as they start at the same time, with probability $(1/2)^{n-1}$ each.

14. Continuous Law of Total Probability?

This has not been covered in class yet, but will be soon.

In this exercise, we will extend the law of total probability to the continuous case.

(a) Suppose we flip a coin with probability U of heads, where U is equally likely to be one of $\Omega_U = \{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$ (notice this set has size $n+1$). Let H be the event that the coin comes up heads. What is $\mathbb{P}(H)$?

Solution:

We can use the law of total probability, conditioning on $U = \frac{k}{n}$ for $k = 0, ..., n$.

$$\mathbb{P}(H) = \sum_{k=0}^{n} \mathbb{P}(H | U = \frac{k}{n})\mathbb{P}(U = \frac{k}{n}) = \sum_{k=0}^{n} \frac{k}{n} \cdot \frac{1}{n+1} = \frac{1}{n(n+1)} \sum_{k=0}^{n} k = \frac{1}{n(n+1)} \frac{n(n+1)}{2} = \frac{1}{2}$$

(b) Now suppose $U \sim$ Uniform(0,1) has the *continuous* uniform distribution over the interval $[0, 1]$. Extend the law of total probability to work for this continuous case. (Hint: you may have an integral in your answer instead of a sum).

Solution:

We can perform basically the same process as above, just using an integral instead of a sum. The values that U can take on are anywhere in the continuous interval $[0, 1]$, so we integrate over that with respect to u. Another change is that we have to use the PDF of U, which in this case is 1 everywhere within our range (since it's uniformly distributed). Plugging that in we can get the same answer of $\frac{1}{2}$ as before.

$$\mathbb{P}(H) = \int_0^1 \mathbb{P}(H|U = u)f_U(u)du = \int_0^1 u \cdot 1du = \frac{1}{2}[u^2]_0^1 = \frac{1}{2}$$

(c) Let's generalize the previous result we just used. Suppose E is an event, and X is a continuous random variable with density function $f_X(x)$. Write an expression for $\mathbb{P}(E)$, conditioning on X.

Solution:

Set up the same problem as before, only this time we're not actually solving for anything. Note that we have to integrate from negative infinity to infinity. We're technically doing this before as well, however outside of the bounds of $[0, 1]$, the density is equal to 0 so the whole expression is equal to 0. In the general case thought, we don't know the range, so we have to integrate everywhere.

$$\mathbb{P}(E) = \int_{-\infty}^{\infty} \mathbb{P}(E|X = x)f_X(x)dx$$

15. Transformations

This has not been covered in class yet and probably won't be. But if you're interested, please read Section 4.4.

Suppose $X \sim \text{Uniform}(0, 1)$ has the continuous uniform distribution on $(0, 1)$. Let $Y = -\frac{1}{\lambda} \log X$ for some $\lambda > 0$.

(a) What is Ω_Y?

Solution:

$\Omega_Y = (0, \infty)$ because $\log(x) \in (-\infty, 0)$ for $x \in (0, 1)$. Thus, that range times a necessarily negative number $-\frac{1}{\lambda}$, will result in a range from 0 to positive infinity.

(b) First write down $F_X(x)$ for $x \in (0, 1)$. Then, find $F_Y(y)$ on Ω_Y.

Solution:

$F_X(x) = x$ for $x \in (0, 1)$ because that is the CDF of the continuous uniform distribution. We find the CDF of Y by plugging in the given definition of Y and getting into a form where we can use the CDF of X. Let $y \in \Omega_Y$.

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(-\frac{1}{\lambda} \log X \le y) = \mathbb{P}(\log X \ge -\lambda y) = \mathbb{P}(X \ge e^{-\lambda y}) = 1 - \mathbb{P}(X < e^{-\lambda y})$$

Then, because $e^{-\lambda y} \in (0, 1)$

$$= 1 - F_X(e^{-\lambda y}) = 1 - e^{-\lambda y}$$

(c) Now find $f_Y(y)$ on Ω_Y (by differentiating $F_Y(y)$ with respect to y. What distribution does Y have?

Solution:

$$f_Y(y) = F'_Y(y) = \lambda e^{-\lambda y}$$

Hence, $Y \sim$ Exponential(λ).

16. Convolutions

This has not been covered in class. We're not yet sure if we will have time for it, but if you're interested, please read Section 5.5.

Suppose $Z = X + Y$, where $X \perp Y$. ($\perp$ is the symbol for independence. In other words, X and Y are independent.) Z is called the convolution of two random variables. If X, Y, Z are discrete,

$$p_Z(z) = \mathbb{P}(X + Y = z) = \sum_x \mathbb{P}(X = x \cap Y = z - x) = \sum_x p_X(x)\, p_Y(z - x)$$

If X, Y, Z are continuous,

$$F_Z(z) = \mathbb{P}(X + Y \le z) = \int_{-\infty}^{\infty} \mathbb{P}(Y \le z - X | X = x) f_X(x) dx = \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx$$

Suppose $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$.

(a) Find an expression for $\mathbb{P}(X_1 < 2X_2)$ using a similar idea to convolution, in terms of $F_{X_1}, F_{X_2}, f_{X_1}, f_{X_2}$. (Your answer will be in the form of a single integral, and requires no calculations – do not evaluate it).

Solution:

We use the continuous version of the "Law of Total Probability" to integrate over all possible values of X_2. Take the probability that $X_1 < 2X_2$ given that value of X_2, times the density of X_2 at that value.

$$\mathbb{P}(X_1 < 2X_2) = \int_{-\infty}^{\infty} \mathbb{P}(X_1 < 2X_2 | X_2 = x_2) f_{X_2}(x_2) dx_2 = \int_{-\infty}^{\infty} F_{X_1}(2x_2)\, f_{X_2}(x_2)\, dx_2$$

(b) Find s, where $\Phi(s) = \mathbb{P}(X_1 < 2X_2)$ using the fact that linear combinations of independent normal random variables are still normal.

Solution:

Let $X_3 = X_1 - 2X_2$, so that $X_3 \sim \mathcal{N}(\mu_1 - 2\mu_2, \sigma_1^2 + 4\sigma_2^2)$ (by the reproductive property of normal distributions)

$$\mathbb{P}(X_1 < 2X_2) = \mathbb{P}(X_1 - 2X_2 < 0) = \mathbb{P}(X_3 < 0) = \mathbb{P}(\frac{X_3 - (\mu_1 - 2\mu_2)}{\sqrt{\sigma_1^2 + 4\sigma_2^2}} < \frac{0 - (\mu_1 - 2\mu_2)}{\sqrt{\sigma_1^2 + 4\sigma_2^2}})$$

$$= \mathbb{P}(Z < \frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}}) = \Phi\left(\frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}}\right) \to s = \frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}}$$