CSE 312: Foundations of Computing II

Section 6: Continuous Random Variables

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 $(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, which has $\mathbb{E}[\overline{X}] = \mu$ and $(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ 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?

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?

5. Max of uniforms

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

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?

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 $\tan 0 = 0$.)

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, is equal to $\frac{\text{A}}{\text{Area of whole circle}}$.
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 $Var(X)$?

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 $Var(X)$.

10. Normal questions

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

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

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

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

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 $\mathbb{P}(X \geq 87)$.

(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.

12. Batteries and exponential distributions

Let X_1, X_2 be independent exponential random variables, where X_i has parameter λ_i, for $1 \leq i \leq 2$. 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?

(b) What is $Pr(X_1 < X_2)$? (Use the law of total probability.)

(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.

[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

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

(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.)

13. 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).

(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).

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 $\mathbb{P}(H)$?

(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).

(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.

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 $X \sim$ 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?

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

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

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 $X \perp Y$. (T$\perp$ 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,

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

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 $F_{X_1}, F_{X_2}, f_{X_1}, f_{X2}$. (Your answer will be in the form of a single integral, and requires no calculations – do not evaluate it).

(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.