CSE 312: Foundations of Computing II

Continuous Random Variables 8 Solutions

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) **Univariate: 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	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

(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) **Multivariate: Discrete to Continuous:**

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X=x, Y=y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X=x, Y=y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x, s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\, ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Conditional PMF/PDF	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X \mid Y=y] = \sum_x x p_{X\|Y}(x\|y)$	$\mathbb{E}[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y) dx$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

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

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

(k) **Law of Total Expectation (Continuous):** Y is a random variable, and X is a continuous random variable with density function $f_X(x)$.

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y \mid X=x]\, f_X(x) dx$$

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:

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

$\mathbb{E}[X] = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$. $F_X(x) = 1 - e^{-\lambda x}$ for $x \ge 0$. 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:

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

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 $\Phi(z) = F_Z(z) = \mathbb{P}(Z \le z)$. Note from symmetry of the probability density function about $z=0$ that: $\Phi(-z) = 1 - \Phi(z)$.

Memorylessness of Exponential

Let $s, t > 0$ be positive real numbers, and $X \sim$ Exponential(λ). Prove the memoryless property of the exponential distribution: $\mathbb{P}(X > s+t | X > s) = \mathbb{P}(X > t)$. (Hint: Use the fact that $\mathbb{P}(X > x) = 1 - F_X(x) = e^{-\lambda x}$). Interpret what this statement is saying. Note that this is also true of the Geometric distribution.

Solution:

$$\mathbb{P}(X > s+t | X > s) = \frac{\mathbb{P}(X > s+t \cap X > s)}{\mathbb{P}(X > s)} = \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t)$$

This means that, after waiting s units of time, the probability you wait t more units of time is the same as waiting t units of time from the beginning.

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:

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

so $c = 2/\pi$.

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

Transformations

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?

Solution:

$\Omega_Y = (0, \infty)$ because $\log(x) \in (-\infty, 0)$ for $x \in (0,1)$.

(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 $x \in (0,1)$. 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. What distribution does Y have?

Solution:

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

Hence, $Y \sim$ Exponential(λ).

Uniform Distribution on the Circle

Consider the closed unit circle of radius r, $S = \{(x, y) : x^2 + y^2 \leq r\}$. Suppose we throw a dart onto this circle and are guaranteed to hit it. However, we would like the probability of hitting any point equally likely. Let (X, Y) be the coordinates of the circle that the dart hits. Find their joint density $f_{X,Y}(x, y)$ and be sure to specify the values for all $(x, y) \in \mathbb{R}^2$. Are X and Y independent? Are the marginal distributions $f_X(x)$ and $f_Y(y)$ uniform on $[-r, r]$?

Solution:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi r^2} & \text{if } (x,y) \in S \\ 0 & \text{otherwise} \end{cases}$$

X and Y cannot be independent since their range fails to be a rectangle ($\Omega_{X,Y} = S \neq [-1,1] \times [-1,1] = \Omega_X \times \Omega_Y$). The marginal distributions are not uniform; they are more likely to be close to the center than the edges.

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:

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

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

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

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

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

For this section, we expect to end here (or before!). The rest of these problems can be done at home for extra practice, or if you finish 2-6 early. Solutions will be posted.

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:

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

Continuous Law of Total Probability?

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:

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

(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:

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

Continuous Law of Total Expectation?

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

(a) Suppose we have fixed a $u \in (0,1)$. We keep drawing independently from the continuous uniform distribution Uniform$(0,1)$ until we draw a number less than u. What is the expected number of times you draw?

Solution:

Let X be the number of draws. Then $X \sim$ Geometric(u) since the probability a uniform rv between 0 and 1 is less than u is exactly u. Hence the expectation is $\frac{1}{u}$.

(b) Suppose we draw from $U \sim$ Uniform$(0,1)$. We keep drawing independently from the continuous uniform distribution Uniform$(0,1)$ until we draw a number less than U. What is the expected number of times you draw? Note that this is different from part (a) since U is random and not fixed. Hint: Consider conditioning on $U = u$ and extend the law of total expectation from the discrete case.

Solution:

$$\mathbb{E}[X] = \int_0^1 \mathbb{E}[X \mid U = u]\, f_U(u)du = \int_0^1 \frac{1}{u}du = \infty$$

(c) Let X be the number of draws in the previous part, and let Y (simultaneously) be the number of draws until the first time we draw a number **greater** than the initial draw (same as X, but X is the number of draws until we get **less** than the initial draw). By symmetry, $\mathbb{E}[X] = \mathbb{E}[Y]$. Let $Z = \min\{X, Y\}$. Explain what Z is in English. What is $\mathbb{E}[Z]$? Is it equal to $\min\{\mathbb{E}[X], \mathbb{E}[Y]\}$?

Solution:

Z is the number of draws until we get a number greater than or less than the initial draw. The first draw after the initial draw is guaranteed to be greater than or less than the first (since with probability 0 they are equal). Hence $\min\{X, Y\} = 1$ with probability 1, and $\mathbb{E}[Z] = 1 \neq \infty = \min\{\mathbb{E}[X], \mathbb{E}[Y]\}$. Interesting right? We have $\mathbb{E}[X] = \mathbb{E}[Y] = \infty$, yet $\min\{X, Y\} = 1$ w.p. 1.

(d) In part (b), it is tempting and intuitive to think that the answer is simply 2, since we expect the first draw to be $\frac{1}{2}$, and hence expect two draws. However, the probability we get between $[0, 0.01]$ is one percent, but the expectation of waiting is 100. The probability we get between $[0, 0.001]$ is extremely small, but the expectation of waiting is much higher. This is why the expectation was infinite. Now, find the flaw in the following argument which defends the intuitive yet incorrect answer. With probability $1/2$, the first draw is less than the initial draw. With probability $1/2$, the second draw is less than the initial draw. And so on. The draws are stated to be independent of each other. Hence $X \sim$ Geometric$(1/2)$ and $\mathbb{E}[X] = 2$.

Solution:

The flaw is the independence: in the standard coin flipping case, knowing whether the first flip is tails doesn't change the probability of the second being heads of tails. Knowing the first hundred flips are tails doesn't change the probability of the next being tails. However, knowing that the first draw is not less than the initial somewhat makes us believe that the initial draw was a bit larger than $1/2$. Knowing that the first hundred draws are not less than the initial makes us believe that the initial draw is very close to 1 rather than its expectation of $1/2$.

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

Solution:

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y \mid X = x] \, f_X(x) dx$$

Convolutions

Suppose $Z = X + Y$, where $X \perp Y$. 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 $F_{X_1}, F_{X_2}, f_{X_1}, f_{X_2}$. (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)

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

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