CSE 312: Foundations of Computing II

Section 6: Continuous Random Variables

1. Review of Main Concepts

(a) **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 range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
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	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Independence must have	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

(b) **Law of Total Probability (r.v. version):** If X is a discrete random variable, then

$$\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A | X = x) p_X(x) \qquad \text{discrete } X$$

(c) **Law of Total Expectation (Event Version)**: Let X be a discrete random variable, and let events $A_1, ..., A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid A_i]\, \mathbb{P}(A_i)$$

(d) **Conditional Expectation**: See table bbelow. Note that linearity of expectation still applies to conditional expectation: $\mathbb{E}[X + Y \mid A] = \mathbb{E}[X \mid A] + \mathbb{E}[Y \mid A]$

(e) **Law of Total Expectation (r.v. Version)**: Suppose X and Y are random variables. Then,

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}[X \mid Y = y]\, p_Y(y) \qquad \text{discrete version.}$$

(f) **Conditional distributions**

	Discrete	**Continuous**
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	$\mathbb{E}[X \mid Y = y] = \sum_x x p_{X\|Y}(x\|y)$	$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y) dx$

(g)

- Law of Total Probability (continuous)

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A | X = x) f_X(x) dx$$

- Law of total expectation (continuous)

$$\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X \mid Y = y]\, f_Y(y) dy$$

Markov's Inequality: Let X be a non-negative random variable, and $\alpha > 0$. Then, $\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[-NoValue-][X]}{\alpha}$.

Chebyshev's Inequality: Suppose Y is a random variable with $\mathbb{E}[-NoValue-][Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$, $\mathbb{P}(|Y - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}$

Chernoff Bound (for the Binomial): Suppose $X \sim \text{Bin}(n, p)$ and $\mu = np$. Then, for any $0 < \delta < 1$

2. Zoo of Continuous Random Variables

(a) **Uniform**: $X \sim \text{Uniform}(a, b)$ 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**: $X \sim \text{Exponential}(\lambda)$ 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.

3. 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?

4. Max of uniforms

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

5. 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.) 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!

(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.

(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.)

6. Continuous joint density I

The joint probability density function of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) & 0 < x < 1,\ 0 < y < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Verify that this is indeed a joint density function.

(b) Compute the marginal density function of X.

(c) Find $Pr(X > Y)$. (Uses the continuous law of total probability which we have not covered in class as of 11/17.)

(d) Find $P(Y > \frac{1}{2} | X < \frac{1}{2})$.

(e) Find $E(X)$.

(f) Find $E(Y)$

7. Continuous joint density II

The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of W and V is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

8. Variance of the geometric distribution

Independent trials each resulting in a success with probability p are successively performed. Let N be the time of the first success. Find the variance of N.

9. 3 points on a line

(This problem uses the continuous law of total probability which has not yet be covered in class.) Three points X_1, X_2, X_3 are selected at random on a line L (continuous independent uniform distributions). What is the probability that X_2 lies between X_1 and X_3?

10. In between

(Covers ideas that have not been covered in class.) Suppose that X_1 and X_2 are discrete uniform random variables in $\{1, \ldots, 2n\}$ (i.e., X_1 and X_2 are equally likely to take any of the values $1, \ldots, 2n$) and let $Y = \min(X_1, X_2)$. What is the conditional pmf $p_{Y|X_1}(y \mid x_1)$ and conditional CDF $F_{Y|X_1}(y \mid x_1)$. What is $E[Y \mid X_1 = x_1]$? (For the definitions of conditional pmf, conditional CDF, see the review at the top of this worksheet.)

11. Tail bounds

Suppose $X \sim$ Binomial$(6, 0.4)$. We will bound $\mathbb{P}(X \geq 4)$ using the tail bounds we've learned, and compare this to the true result.

(a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?

(b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound.

(c) Give an upper bound for this probability using the Chernoff bound.

(d) Give the exact probability.