CSE 312: Foundations of Computing II

Section 7: Joint Distributions, Law of Total Expectation (and bit of conditional distributions)

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 (RV 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 (not realy covered in class)**

	Discrete	**Continuous**
Conditional PMF/PDF	$p_{X\mid Y}(x\mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\mid Y}(x\mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X \mid Y = y] = \sum_x x p_{X\mid Y}(x\mid y)$	$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X\mid Y}(x\mid y) dx$

(g) **The following have not been covered as of 11/17:**

- 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)

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

2. Joint PMF's

Suppose X and Y have the following joint PMF:

X/Y	1	2	3
0	0	0.2	0.1
1	0.3	0	0.4

(a) Identify the range of X (Ω_X), the range of Y (Ω_Y), and their joint range ($\Omega_{X,Y}$).

(b) Find the marginal PMF for X, $p_X(x)$ for $x \in \Omega_X$.

(c) Find the marginal PMF for Y, $p_Y(y)$ for $y \in \Omega_Y$.

(d) Are X and Y independent? Why or why not?

(e) Find $\mathbb{E}\left[X^3 Y\right]$.

3. Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of n independent trials, but with three outcomes, where $\mathbb{P}(\text{outcome } i) = p_i$ for $i = 1, 2, 3$ and of course $p_1 + p_2 + p_3 = 1$. Let X_i be the number of times outcome i occurred for $i = 1, 2, 3$, where $X_1 + X_2 + X_3 = n$. Find the joint PMF $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$ and specify its value for all $x_1, x_2, x_3 \in \mathbb{R}$.

4. Do You "Urn" to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

(a) X_1, X_2

(b) X_1, X_2, X_3

5. Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability p. Let X_1 be the number of failures preceding the first success, and let X_2 be the number of failures between the first 2 successes. Find the joint pmf of X_1 and X_2. Write an expression for $E[\sqrt{X_1 X_2}]$. You can leave your answer in the form of a sum.

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. Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1: The 1^{st} door leads to a tunnel that will take him to safety after 3 hours.
- D_2: The 2^{nd} door leads to a tunnel that returns him to the mine after 5 hours.
- D_3: The 3^{rd} door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

9. Elevator

[We have done this problem in class.] The number of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where the others get off, compute the expected number of stops that the elevator will make before discharging all the passengers. Assume an infinitely large elevator.

10. Lemonade Stand

Suppose I run a lemonade stand, which costs me $100 a day to operate. I sell a drink of lemonade for $20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1. If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2. It rains each day with probability p_3, independently of every other day. Let X be my profit over the next week. In terms of n_1, n_2, p_1, p_2 and p_3, what is $\mathbb{E}[X]$?

11. Particle Emissions

Suppose we are measuring particle emissions, and the number of particles emitted follows a Poisson distribution with parameter λ, $X \sim \text{Poisson}(\lambda)$. Suppose our device to measure emissions is not always entirely accurate sometimes we fail to observe particles that actually emitted. So for each particle actually emitted, say we have probability p of actually recording it, independently of other particles. Let Y be the number of particles we observed. What distribution does Y follow with what parameters, and what is $\mathbb{E}[Y]$?

12. 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.

13. 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?