**Math 416   Lecture 2**

**Definition.** Here are the multivariate versions:

**PMF case:** \(p(x, y, z)\) is the *joint Probability Mass Function of \(X, Y, Z*\) iff

\[ P(X=x, Y=y, Z=z) = p(x, y, z) \]

**PDF case:** \(f(x, y, z)\) is the *joint Probability Density Function of \(X, Y, Z*\) iff for all sets \(A, B, C\),

\[ P(x \in A, y \in B, z \in C) = \int_A \int_B \int_C f(x, y, z) dz \, dy \, dx. \]

If you know \(p(x, y, z)\) or \(f(x, y, z)\) then you can calculate the distribution of \(X_1\) (or \(X_2\) or \(X_3\)) alone. This distribution of \(X_1\) alone is the *marginal distribution of \(X_1*\).

**PMF case:** \(P(X_1=x_1) = \sum_y \sum_z p(x, y, z)\) summed over all \(y, z\).

**PDF case:** \(P(X=x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dy dz\).

Suppose \(X\) and \(Y\) are random variables such that \(X\) has two values: 0, 1 and \(Y\) has two values 2, 3. Suppose the joint PMF \(p(x, y)\) is \(p(0,2)=2/5, p(0,3)=p(1,2)=p(1,3)=1/5\).

Find \(P(X=1)\).

\[ P(X=1) = P(X=1, Y=2) + P(X=1, Y=3) \]
\[ = p(1,2) + p(1,3) = 1/5 + 1/5 = 2/5. \]
The probability of an event $A$ is the proportion of outcomes in the sample space $\Omega$ which are in $A$. $A$ has probability .5 iff it has 50% of the outcomes.

Suppose an event $A$ of nonzero probability is known to have occurred. Then outcomes not in $A$ are ruled out and $A$ becomes the new sample space. Since $A$ is 100% of this new sample space, $A$ now has probability 1. The conditional probability $P[B|A]$ of another event $B$ is the proportion of outcomes in $A$ which lie in $B$.

**Definition.** The conditional probability of $B$ given $A$ is $P[B|A] = P[A \cap B]/P[A] = \text{the percentage of } A \text{ which lies in } B$.

A rider just misses a bus. The time of arrival of the next bus has an exponential distribution with rate $\lambda = 3 / \text{hour}$.

(a) Find the probability the next bus arrives within $t$ minutes.

(b) Find the probability it arrives later than $t$ minutes.

(c) If the rider has waited $s$ minutes, find the probability the bus will take $t$ minutes or more before it arrives.
A rider just misses a bus. The time of arrival of the next bus has an exponential distribution with rate \( \lambda = 3 \) / hour.

(a) Find the probability the next bus arrives within \( t \) minutes.
\[
F(t) = 1 - e^{-3t}
\]

(b) Find the probability it arrives later than \( t \) minutes.
\[
1 - F(t) = 1 - (1 - e^{-3t}) = e^{-3t}
\]

(c) If the rider has waited \( s \) minutes, find the probability the bus will take \( t \) minutes or more before it arrives.
\[
e^{-3(s+t)}/e^{-3s} = e^{-3t}
\]
**Definition.** The *conditional probability* of $A$ given $B$ is $P[A|B] = P[B \cap A]/P[B]$.


**Definition.** Events $A$ and $B$ are *independent* iff $P[A \cap B] = P[A]P[B]$.

**Bayes Theorem.** Suppose events $B_1, B_2, B_3$ partition the sample space, i.e., they are disjoint and their union is everything. Then for any event $A$,


$$= \sum_i P[A|B_i]P[B_i].$$

**Proof.** $P[A] = P[A \cap B_1] + P[A \cap B_2] + P[A \cap B_3]$


This is easy to visualize with a Venn diagram.

![Venn Diagram](image)

The formulas for conditional probability mass and density functions obey the same laws.
Suppose $X$ and $Y$ are two continuous random variables on the same sample space, say the height and weight respectively of a randomly picked UH student. Let $f(x, y)$ be the joint density function. Hence $P[\text{height is from 5’ to 6’ & weight is from 100 lb to 150 lb}] = \int_5^6 \int_{100}^{150} f(x, y) dx\,dy$.

Let $f_X(x)$ be the (marginal) density function for $X$. This is the probability density function for $x$. We don’t care what $y$ is, it can be anything. Thus to get the marginal density we integrate over all possible values of $y$ --

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$ This is the cross-section at $x$.

Since $f(x, y)$ is the joint distribution of $X$ and $Y$ and since $\cap$ is the set theoretic translation of “and”, the conditional formula $P[Y|X] = P[X \cap Y]/P[X]$ translates into the following formula for density functions.

**Definition.** The *conditional density* of $Y$ given $X=x$ is $f(y|x) = f(x, y)/f_X(x)$.

For any set $B$, $P[Y \in B|X = x] = \int_{y \in B} f(y|x) dy$. 
More generally, if \( f(x_1, x_2, y_1, y_2) \) is the joint density of random variables \( X_1, X_2, Y_1, Y_2 \), then the conditional density of \( Y_1, Y_2 \) given \( X_1 = x_1, X_2 = x_2 \) is

\[
f(y_1, y_2 | x_1, x_2) = \frac{f(x_1, x_2, y_1, y_2)}{f_{X_1, X_2}(x_1, x_2)}
\]

where \( f_{X_1, X_2}(x_1, x_2) = \int f(x_1, x_2, y_1, y_2) dy_1 dy_2 \) is the marginal distribution of \( X_1, X_2 \).

Recall: events (sets) \( A, B, C \) were independent if

\[
P[A \cap B \cap C] = P[A] P[B] P[C].
\]

Suppose \( X, Y, Z \) are random variables with joint probability density function \( f(x, y, z) \).

**Definition.** \( X, Y, Z \) are mutually independent iff

\[
P[X \in A, Y \in B, Z \in C] = P[X \in A] P[Y \in B] P[Z \in C]
\]

for all sets \( A, B, C \). Equivalently, \( X, Y, Z \) are mutually independent iff

\[
f(x, y, z) = f_X(x) f_Y(y) f_Z(z)
\]

where \( f_X(x), f_Y(y), f_Z(z) \) are their marginal density functions.

- Suppose \( X, Y \) have joint density function \( f(x, y) = 1/6 \) if \( x \in [0, 2], y \in [0, 3], = 0 \) elsewhere.

  Are \( X \) and \( Y \) independent?

  Find \( f_X(x), f_Y(y) \). Is \( f(x, y) = f_X(x) f_Y(y) \)?
Are height and weight independent random variables?

You randomly pick 3 students. Let $X$, $Y$, $Z$ be the heights of the first, second and third student.

Are $X$, $Y$, and $Z$ independent?

Do they have the same marginal distribution, i.e., is $f_X(x) = f_Y(y) = f_Z(z)$?

As is usually the case in statistics, they have the normal distribution. The density function is the normal bell-curve

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

where $\mu$ is the average and $\sigma$ is the standard deviation.
Suppose \( f(x, y) \) is the joint density function of random variables \( X \) and \( Y \) then

\[
f(y|x) = \frac{f(x, y)}{f_X(x)} \text{ is the conditional density of } Y \text{ and}
\]

\[
f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy \text{ is the marginal density for } X.
\]

\[
f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx \text{ is the marginal density for } Y.
\]

Density analog of Bayes Theorem: \( P[B] = \sum_i P[B|A_i]P[A_i] \)

**Continuous Bayes Theorem.**

\[
f_Y(y) = \int_{-\infty}^{\infty} f(y|x)f_X(x)dx.
\]

**Proof.**

\[
f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{f_X(x)} f_X(x)dx
\]

\[
= \int_{-\infty}^{\infty} f(y|x) f_X(x)dx.
\]

**Conditioning/unconditioning Theorem.**

\[
\int_A \int_B f(x, y)dydx = \int_A (\int_B f(y|x)dy)f_X(x)dx
\]

**Proof.**

\[
\int_A \int_B f(x, y)dydx
\]

\[
= \int_A \int_B \frac{f(x,y)}{f_X(x)} f_X(x)dydx
\]

\[
= \int_A (\int_B \frac{f(x,y)}{f_X(x)} dy)f_X(x)dx
\]

\[
= \int_A (\int_B f(y|x)dy)f_X(x)dx.
\]

\[
= \int_{-\infty}^{\infty} (\int_B f(y|x)dy)f_X(x)dx.
\]