# MATH 3338 Probability

Wenjiang Fu
wfu2@central.uh.edu

Lecture 2 - 3338

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Outline


1. Sets and Probability
2. Continuous Probability Densities
3. Continuous Probability Density

# Sets

- An **event** is a set of elements in the sample space $\Omega$, it is a subset of the sample space.
- Given two or more sets in the sample space $\Omega$, we learned their operations, joint (union), intersection, complement, etc.
- Q: How to calculate the probability of these sets in operation?
- There are probability rules or laws that we need to follow to calculate the probability of these operation of sets.

# Properties of Probability

- **Properties**

  **Theorem 1.1** The probabilities assigned to events by a distribution function on a sample space $\Omega$ satisfy the following properties:

  1. $P(E) \geq 0$ for every $E \subset \Omega$.
  2. $P(\Omega) = 1$.
  3. If $E \subset F \subset \Omega$, then $P(E) \leq P(F)$.
  4. If $A$ and $B$ are disjoint subsets of $\Omega$, then $P(A \cup B) = P(A) + P(B)$.
  5. $P(\bar{A}) = 1 - P(A)$ for every $A \subset \Omega$.

- **Proof** The proof uses the following basics.

  For any event $E$ the probability $P(E)$ is determined from the distribution $m$ by

$$P(E) = \sum_{\omega \in E} m(\omega),$$

for every $E \subset \Omega$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Properties**

**Theorem 1.2** If $A_1, ..., A_n$ are pairwise disjoint subsets of $\Omega$ (i.e. no two of the $A_i$'s have an element in common), Then

$$P(A_1 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i).$$

**Theorem 1.3** Let $A_1, ..., A_n$ be disjoint events with $\Omega = A_1 \cup ... \cup A_n$, and let $E$ be any event. Then

$$P(E) = \sum_{i=1}^{n} P(E \cap A_i)$$

Theorem 1.3 will be extended to a total probability law later with Bayes Theorem. The total probability law has interesting stories associated with it. One is called a Simpson's paradox, which also got involved in a court ruling.

# Properties of Probability

- **Properties**
  **Corollary 1.1** For any two events $A$ and $B$,

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

**Theorem 1.4** If $A$ and $B$ are two subsets of $\Omega$, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- There are a number of interesting paradox examples in our textbook. For those who would like to read them with fun, happy reading.

- Probability is fun and is also intriguing.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Example of Continuous Probability Density

Example 1. Throw a dart on a unit circle. How likely you may hit the small circle of one half unit.
How to calculate?
What kind of assumption(Bertrand's paradox) do you make?
Calculation based on the assumption?

Example 2. (Bertrand's paradox) Randomly draw a chord connecting two points on a unit circle. What's the probability of having a chord with length greater than $\sqrt{3}$?
Q: Why is it a paradox?
Depends on how you do this *randomly*.
a) Randomly choose points on the unit circle.
b) Choose the point on the circle by moving a radius of the circle from the radius connecting the center of the circle and a given point on the circle by a random angle.

# Continuous Random Variable

- In the above example, the random points are taking “values” in a continuum, like the points on a circle, or a disk, or even a rectangle, a 3D cube, etc.

- How to specify the probability in such a sample space of continuum?

# Analogy in Discrete Case

| | | | | | |
|---|---|---|---|---|---|
| (1,1) | (2,1) | (3,1) | (4,1) | (5,1) | (6,1) |
| (1,2) | (2,2) | (3,2) | (4,2) | (5,2) | (6,2) |
| (1,3) | (2,3) | (3,3) | (4,3) | (5,3) | (6,3) |
| (1,4) | (2,4) | (3,4) | (4,4) | (5,4) | (6,4) |
| (1,5) | (2,5) | (3,5) | (4,5) | (5,5) | (6,5) |
| (1,6) | (2,6) | (3,6) | (4,6) | (5,6) | (6,6) |

Q: How do we specify the probability ?

Calculate the probability of each small cell.

If the number of squares (cell) goes larger and larger, how to do it?

# Analogy in Discrete Case

- Specify a density of each square (cell) instead of specifying a probability.
- Calculate the density by dividing the probability of each square by its area.
- Equivalently to say, the probability = density multiplied by its area.
- For the limiting case, when the squares (cells) are so small, the probability of each piece goes to 0, so specify a density for each.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Density function

- For a given domain of a continuous random variable $X$, say, in a 1-dim space, like, [0, 1]. Make a density function $f(x)$ by

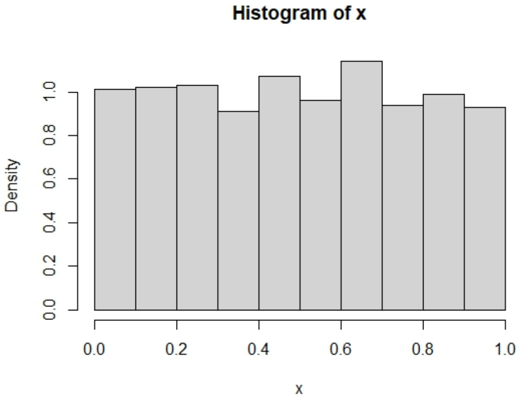
$$f(x) = 1 \ \ \forall x \in [0, 1]$$

  It's called uniform distribution on $[0, 1]$.

- Of course, the above example can be extended to many different continuous sample spaces, such as unit disk, unit circle, or the entire 2-dim space.

# Continuous Probability Examples

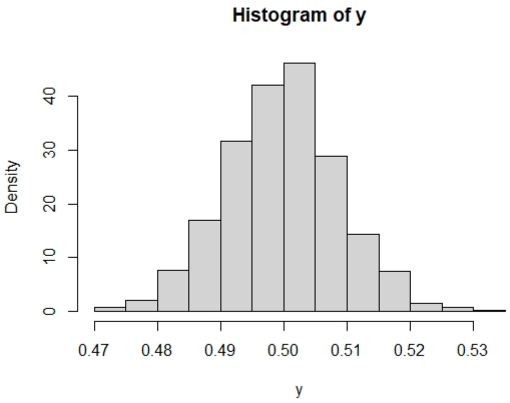
- Take 1000 random variates (random numbers) from the above uniform distribution $[0, 1]$. Look at the distribution by plotting the histogram.




```
x = runif(1000,0,1); hist(x); hist(x, prob=T)
```

# Continuous Probability Examples

- If we take the mean of above 1000 random variates from [0, 1] repeatedly and plot 1000 means.




```
y = rep(NA, 1000); for (i in 1:1000) y[i] =
mean(runif(1000, 0, 1)); hist(y, prob=T)
```

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Density function

- From the above, it seems that it is more appropriate to specify a density function $f(x)$ for a continuous random variable $X$, so that the behavior of $X$ will follow the density for the probability distribution.

- **Note** A density function $f(x) \geq 0$.
  Density function can be considered for two or more continuous random variables.

- **Definition 2.1** Let $X$ be a continuous real-valued random variable. A density function for $X$ is a real-valued function $f$ which satisfies

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

  for all $a, b \in \mathbb{R}$. The Probability is equal to the area under the density function.

# Density function

- Examples of density functions
  Example 1. Uniform distribution on interval $[a, b]$ with $a < b$.
  By the definition, a uniform distribution over an interval $[a, b]$ will have a constant density so that any interval has the same probability over a sub-interval of the same length.
  Assume the density $f(x) = C$, a constant. Then by the definition of density function,

$$1 = P(a \leq X \leq b) = \int_a^b f(x)dx = \int_a^b Cdx = C(b - a)$$

which implies that $C = \frac{1}{b-a}$. i.e. $f(x) = \frac{1}{b-a}$.

# Density function

- Examples of density functions
  Example 2. A density function $f(x) = 2x$ on $[0, 1]$.
  The total probability is

$$P(0 \leq X \leq 1) = \int_0^1 f(x)dx = \int_0^1 2xdx = 1$$

Probability over an interval $[a, b]$ with $0 \leq a \leq b \leq 1$

$$P(a \leq X \leq b) = \int_a^b 2xdx = b^2 - a^2$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Density function

- Examples of density functions
  Example 3 (2.11). A density function of two random variables $X, Y$ is defined as $f(x, y) = 1/\pi$ on the unit disk $D = \{(x, y)|x^2 + y^2 \leq 1\}$.
  The total probability is

$$P(x^2 + y^2 \leq 1) = \int\int_D f(x)dx = \int\int_D 1/\pi dx = \pi/\pi = 1$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Cumulative Distribution Function

- In addition to the probability density function (PDF) for continuous random variables, the cumulative distribution function (CDF) is also important to provide a way to assess the probability.
- The cumulative distribution function for a random variable $X$ is defined as $F(x) = \int_{-\infty}^{x} f(x)dx$ for any given $x \in \mathbb{R}$
  The CDF $F(x)$ is an increasing and right continuous function.

$$0 \le F(x) \le 1, \quad \text{and} \quad \frac{d}{dx}F(x) = f(x).$$

Examples of CDF.
Uniform distribution.

$$F(x) = \int_{a}^{x} C dx = C(x-a) = (x-a)/(b-a), \quad \forall x \in [a,b],$$

Linear density function $f(x) = 2x$ for $x \in [0,1]$.

$$F(x) = \int_{0}^{x} 2x dx = x^2, \quad \forall x \in [0,1].$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Cumulative Distribution Function

- **Example 1** of CDF (2.13).
  A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is squared. Let $Y$ represent the result. What is the cumulative distribution function of $Y$? What is the density of $Y$?
  **Solution** Let $X \sim Unif[0, 1]$, $Y = X^2$. To figure out the pdf and cdf of $Y$, we find the cdf from the relationship between $X$ and $Y$.

$$F_y(y) = Pr(Y \leq y) = Pr(X^2 \leq y) = Pr(X \leq \sqrt{y}), \ \forall y \geq 0$$

$$= F_x(x)|_{x=\sqrt{y}} = x|_{x=\sqrt{y}} = \sqrt{y}$$

  Finally, summarize the results as follows. $F_Y(y) = \sqrt{y} \ \forall y \in [0, 1]$, 0 for $Y \leq 0$, and 1 for $Y \geq 1$. And the density of $Y$ is $f_Y(y) = 1/(2\sqrt{y})$ for $y \in [0, 1]$ and 0 otherwise.
  $F(y)$ is continuous, but $f(y)$ is not.

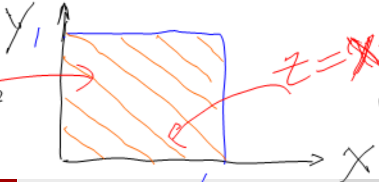
UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Example 2** of CDF (2.14).
  In example 2.4 of the textbook, the sum of two random variables from *Unif*$[0, 1]$ is considered, now consider the distribution of their sum. Denote by $X, Y$ the two random variables from *Unif*$[0, 1]$. Find the pdf and cdf of $Z = X + Y$.
  **Attention!**. In the book, the authors mentioned that a random experiment was done for the sum $X + Y$. It implies that the two RV $X$ and $Y$ are independent. That means the outcome of one does not affect the other.
  If not true, they are called correlated. Extreme case is $X = Y$. Their sum would not follow a "bell-curve" distribution.
  **Q:** What is the distribution of $X + Y$ now for $X = Y$?
  **A:** The distribution of $Y = X$, $Z = X + Y = 2X$, uniformly distributed on $[0, 2]$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Example 2** of CDF (2.14) assuming independence.
  **Q:** What can we start to consider this problem with? PDF or CDF ?
  Often, one may use the relationship between variables and work on the CDF (prob). For this one,
  $F_Z(z) = Pr(X + Y \leq z)\ \forall z \in [0, 2]$ ?
  But for this one, a very special distribution of $X$ and $Y$. We may want to start with the pdf. Why?
  To make it easy to work on.
  **Q:** Can we guess what kind of distribution we have for $X + Y$ if they are indep.

$(X, Y)$ unif $[0,1]^2$




$z = \cancel{X}+Y$

$f(z)$ linear in $z$

# Cumulative Distribution Function

- **Example 2** of CDF (2.14) assuming independence.
  Since $(X, Y)$ are jointly distributed uniformly on the unit square $[0, 1] \times [0, 1]$, just need to figure out the distribution $Z = X + Y$ on $[0, 2]$. As a function of $(X, Y)$, it is uniformly dist on the unique square. Correct?
  Look at the unit square, $Z = X + Y$ on the diagonals. Consider the prob. Is the segment length of $Z$ increases with $X$ or $Y$ linearly in the square? then the area of trapezoid in a strap $(z, z + dz)$ ? linearly depend on $z$? So the density (corresponding to the area with an increment of $dz$), is $f(z) = az$ (through 0) for $0 \leq z \leq 1$. After that the symmetry applies for $z \in [1, 2]$: $f(z) = b - az$. Use prob to solve for constants. Take the integral for $F(z)$.

$$z \in [0,1] : F(z) = c_1 + az^2/2; \;\; z \in [1,2] : F(z) = c_2 - az^2/2 + bz$$

$$0 = F(0) \Rightarrow c_1 = 0, \;\; 1 = F(2) \Rightarrow c_2 - 2a + 2b = 1$$

$$F(1) = 0.5 \Rightarrow 0.5 = c_2 - a/2 + b = a/2 \Rightarrow a = 1, b = 2, c_2 = -1$$

# Cumulative Distribution Function

- **Example 2** of CDF (2.14) assuming independence.
  In summary, under the independence assumption of $X$ and $Y$ following $Unif[0,1]$, we have the pdf and cdf for $Z = X + Y$.

  $z \in [0,1] : f(z) = z; \;\; z \in [1,2] : f(z) = 2 - z; \;\; z \notin [0,1] : f(z) = 0.$

$$f(z) = \begin{cases} 0, & z < 0, \\ z, & 0 \leq z \leq 1 \\ 2 - z, & 1 \leq z \leq 2 \\ 0, & z > 2. \end{cases}$$