# Math 425
# Introduction to Probability
# Lecture 20

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

March 4, 2009





## Cumulative Distribution Function

☞ There is another way to give the probability of a continuous random variable, that is often easier to find then the density function.

**Definition**

Let $X$ be a continuous real-valued random variable. The cumulative distribution function of $X$ is defined by the equation

$$F_X(a) = \mathbf{P}\{X \le a\} \qquad \text{for all } a \in \mathbb{R}.$$





## Example

**Example.** Let $X$ be a uniform random variable on the interval $[0, 2\pi)$. What is the cumulative distribution for $X$?

☞ The density $f_X$ is given by

$$f_X(t) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \le t < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

The cumulative distribution $F_X$ is given by

$$F_X(a) = \int_{-\infty}^{a} \frac{1}{2\pi}\, dt = \begin{cases} 0 & \text{if } a < 0 \\ \frac{a}{2\pi} & \text{if } 0 \le a \le 2\pi \\ 1 & \text{otherwise} \end{cases}$$

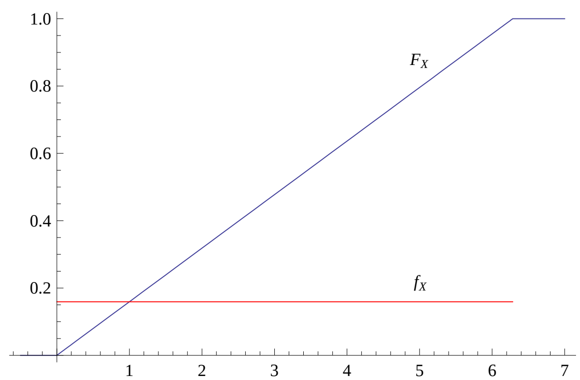




## Example – continued

☞ The cumulative distribution and density for a uniform random variable on $[0, 2\pi)$.

Note that $F_X$ is continuous, although $f_X$ is not.



## Cumulative Distribution and Basic Events

☞ Let $X$ be a continuous random variable with cumulative distribution $F_X$. Then for any real numbers $a$ and $b$

$$\begin{aligned} 1 &= \mathbf{P}\{a \le X \le b\} + \mathbf{P}\{X \le a\} + \mathbf{P}\{b < X\} \\ &= \mathbf{P}\{a \le X \le b\} + \mathbf{P}\{X \le a\} + (1 - \mathbf{P}\{X \le b\}) \\ &= \mathbf{P}\{a \le X \le b\} + F_X(a) + (1 - F_X(b)). \end{aligned}$$

So, for all real numbers $a$ and $b$

$$\mathbf{P}\{a \le X \le b\} = F_X(b) - F_X(a).$$

or equivalently, when $f_X$ is the density of $X$,

$$\int_a^b f_X(t)\,dt = F_X(b) - F_X(a).$$





## Theorem

☞ The density function $f_X$ and cumulative distribution function $F_X$ of a continuous random variable $X$ are related in a very nice way. The following is just the statement of the Second Fundamental Theorem of Calculus.

### Theorem

*Suppose $X$ is a continuous random variable with density $f_X(t)$ which is continuous, except perhaps finitely many points.*

*The cumulative distribution function $F_X$ is given by*

$$F_X(a) = \int_{-\infty}^a f_X(t)\,dt \qquad \text{for each } a \in \mathbb{R}$$

*It is continuous and is related to $f_X$ by*

$$\frac{d}{da}F_X(a) = f_X(a).$$





## Theorem

☞ This gives a converse of the previous theorem. Given the cumulative distribution function $F_X$, you can recover the density function $f_X$. It is really just a restatement of the First Fundamental Theorem of Calculus.

### Theorem

*Suppose $X$ is a continuous random variable with cumulative distribution function $F_X$.*

*The density function $f_X$ is given by*

$$\frac{d}{dt}F_X(a) = f_X(a).$$

*and is related to $F_X$ by*

$$F_X(a) = \int_{-\infty}^a f_X(t)\,dt \qquad \text{for each } a \in \mathbb{R}$$





## Example – return

☞ Sometimes it is easier to specify the cumulative distribution function, then it is the density function.

**Example.** A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is squared. $X$ is the r.v. which represents this is result.

What is the cumulative distribution of $X$? What is the density of $X$?

☞ Let $U$ be the uniform r.v. giving the chosen number. So, $X = U^2$.

The key is that we know the density of $U$ already.

# Example – continued

☞ Let $U$ represent the chosen real number and $X = U^2$.
For $0 \leq a \leq 1$,

$$\begin{aligned} F_X(a) &= \mathbf{P}\{X \leq a\} \\ &= \mathbf{P}\{U^2 \leq a\} \\ &= \mathbf{P}\{U \leq \sqrt{a}\} \\ &= \int_0^{\sqrt{a}} dt = \sqrt{a}. \end{aligned}$$

Since $0 \leq X \leq 1$, the cumulative distribution function for $X$ is

$$F_X(a) = \begin{cases} 0 & \text{if } a \leq 0, \\ \sqrt{a} & \text{if } 0 \leq a \leq 1, \\ 1 & \text{if } a \geq 1. \end{cases}$$

The density function of $X$ is obtained by differentiating $F_X$

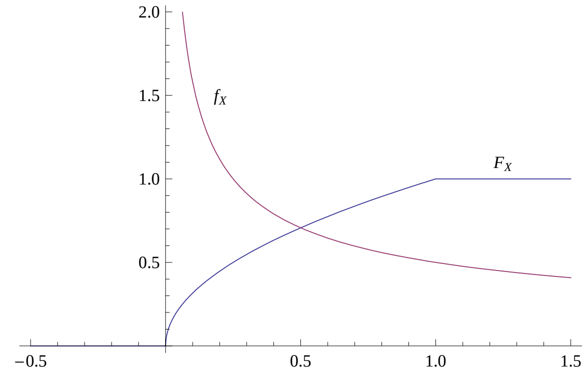
$$f_X(a) = \begin{cases} \frac{1}{2\sqrt{a}} & \text{if } 0 \leq a \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$





# Example – continued

☞ The cumulative distribution $F_X(t) = \sqrt{t}$ and density function $f_X(t) = \frac{1}{2\sqrt{t}}$ on $[0, 1]$. Note that $F_X$ is continuous, although $f_X$ is not.



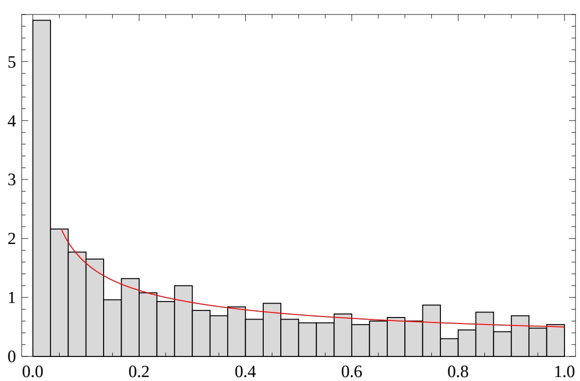





# Example – continued

☞ A plot of 1000 randomly chosen and squared numbers in the interval $[0, 1]$ and partitioned into 30 equal subintervals.

The red line is $f(t) = \frac{1}{2\sqrt{t}}$, and is very close to giving the area of the rectangles in each subinterval. Note that $f(t)$ blows-up at 0.








# Example – return

**Example.** A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is halved. $X$ is the r.v. which represents this is result.

What is the cumulative distribution of $X$? What is the density of $X$?

☞ Let $U$ be the uniform r.v. giving the chosen number. So, $X = \frac{U}{2}$.
We know the density of $U$ already.

## Example – continued

- Let $U$ represent the chosen real number and $X = \frac{U}{2}$.
  For $0 \le a \le \frac{1}{2}$ (the possible values of $X$)

$$\begin{aligned} F_X(a) &= \mathbf{P}\{X \le a\} \\ &= \mathbf{P}\{\frac{U}{2} \le a\} = \mathbf{P}\{U \le 2a\} \\ &= \int_0^{2a} dt = 2a. \end{aligned}$$

Since $0 \le X \le \frac{1}{2}$, the cumulative distribution function for $X$ is

$$F_X(a) = \begin{cases} 0 & \text{if } a \le 0, \\ 2a & \text{if } 0 \le a \le \frac{1}{2}, \\ 1 & \text{if } a \ge \frac{1}{2}. \end{cases}$$

The density function of $X$ is obtained by differentiating $F_X$

$$f_X(a) = \begin{cases} 2 & \text{if } 0 \le a \le \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$





## Expectation

### Definition

Let $X$ be a random variable with density function $f(t)$. The expected value $\mu = E[X]$ is defined by

$$\mu = E[X] = \int_{-\infty}^{+\infty} t\, f(t)\, dt,$$

provided the integral

$$\int_{-\infty}^{+\infty} |t|\, f(t)\, dt,$$

is finite.





## Expectation of a function of a r.v.

### Theorem

*Let $X$ be a random variable with distribution $f_X$ and $g : \mathbb{R} \to \mathbb{R}$ a continuous function (except, perhaps on finitely many points). Then*

$$E[g(X)] = \int_{-\infty}^{+\infty} g(t) f_X(t)\, dt$$

*provided this integral exists.*

- ☠ Warning. It is not generally true that $E[g(X)] = g(E[X])$.
  For example, it is not true that $E[X^2] = \left(E[X]\right)^2$. An example to follow.





## Corollary

### Corollary

*Let $a$ and $b$ be real numbers. Then*

$$E[aX + b] = aE[X] + b$$

### Proof.

Let $Y = aX + b$ and $f_X$ be the density of $X$. Then

$$\begin{aligned} E[Y] &= E[aX + b] = \int_{-\infty}^{+\infty} (at + b)\, f_X(t)\, dt \\ &= a \int_{-\infty}^{+\infty} t\, f_X(t)\, dt + \int_{-\infty}^{+\infty} b\, f_X(t)\, dt \\ &= aE[X] + b. \end{aligned}$$

□

## Variance

**Definition**

The variance for a continuous random variable is defined exactly as for discrete random variables: let $\mu = E[X]$, then

$$Var(X) = E[(X-\mu)^2].$$

☞ Just as with variance for discrete random variables,

$$Var(X) = E[X^2] - \big(E[X]\big)^2$$

☞ Just like with discrete r.v. (the same argument works)

$$Var(aX+b) = a^2 Var(X).$$





## Example – Expectation

**Example.** Let $U$ be a uniformly distributed on the interval $[\alpha, \beta]$. Since the density $f_U(t) = \frac{1}{\beta-\alpha}$ on $[\alpha, \beta]$ and 0 elsewhere:

$$\begin{aligned} E[U] &= \int_\alpha^\beta t \frac{1}{\beta-\alpha}\, dt \\ &= \frac{t^2}{2(\beta-\alpha)}\Big|_\alpha^\beta \\ &= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} \\ &= \frac{\beta+\alpha}{2} \end{aligned}$$

☞ The expectation of a uniformly distributed r.v. on the interval $[\alpha, \beta]$ is the midpoint of the interval $\frac{\beta+\alpha}{2}$.





## Example – Variance

**Example.** Let $U$ be a uniformly distributed on the interval $[\alpha, \beta]$. Since $E[U] = \frac{\alpha+\beta}{2}$ (the midpoint),

$$\begin{aligned} E[U^2] &= \int_\alpha^\beta \frac{t^2}{\beta-\alpha}\, dt \\ &= \frac{t^3}{3(\beta-\alpha)}\Big|_\alpha^\beta \\ &= \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} = \frac{\beta^2+\alpha\beta+\alpha^2}{3} \\ Var(U) &= E[U^2] - \big(E[U]\big)^2 = \frac{\beta^2+\alpha\beta+\alpha^2}{3} - \frac{(\alpha+\beta)^2}{4} \\ &= \frac{(\beta-\alpha)^2}{12} \end{aligned}$$

For example, if $U$ is uniformly distributed on $[0, 1]$,

$$E[U] = \frac{1}{2} \qquad Var(U) = \frac{1}{12} \qquad SD(U) = \sqrt{Var(U)} \approx 0.289$$





## Example – Variance

☞ Let $U$ be a uniformly distributed on the interval $[\alpha, \beta]$. Then

$$\begin{aligned} E[U^2] &= \frac{\beta^2+\alpha\beta+\alpha^2}{3} \\ \big(E[U]\big)^2 &= \Big(\frac{\alpha+\beta}{2}\Big)^2 \\ &= \frac{(\alpha+\beta)^2}{4}. \end{aligned}$$

$E[U^2] \neq \big(E[U]\big)^2$, unless $\alpha = \beta = 0$.
For example, if $U$ is uniformly distributed on $[0, 1]$,

$$E[U^2] = \frac{1}{3} \qquad \big(E[U]\big)^2 = \frac{1}{4}.$$

# Example

## Example

A number in the interval $[0, 1]$ is chosen at random and multiplied by $a$. What is the expected value and variance of the number?

☞ Let $U$ be uniformly distributed on $[0, 1]$.

$$
\begin{aligned}
E[aU] &= aE[U] = \frac{a}{2} \\
E[(aU)^2] &= E[a^2U^2] = a^2E[U^2] = \frac{a^2}{3} \\
Var(aU) &= E[(aU)^2] - \big(E[aU]\big)^2 \\
&= \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12} \\
&= a^2 Var(U) \qquad Var(U) = \frac{1}{12}.
\end{aligned}
$$





# Example

## Example

Let $U$ be a uniformly distributed r.v. on $[0, 1]$. What is the expectation and variance of $U^2$?

☞ I'll first compute the expectation directly from the density $f_{U^2}$. We saw earlier the density of $U^2$ is

$$
f_{U^2}(t) = \begin{cases} \frac{1}{2\sqrt{t}} & \text{if } t \in [0, 1] \\ 0 & \text{otherwise.} \end{cases}
$$

So,

$$
E[U^2] = \int_0^1 t \frac{dt}{2\sqrt{t}} = \int_0^1 \frac{t^{\frac{1}{2}}}{2}\, dt = \frac{t^{\frac{3}{2}}}{3}\Big|_0^1 = \frac{1}{3}
$$





# Example – continued

☞ I recompute expectation from the density $f_U$.

$$
\begin{aligned}
E[U^2] &= \int_0^1 t^2 f_U(t)\, dt \\
&= \int_0^1 t^2\, dt = \frac{t^3}{3}\Big|_0^1 = \frac{1}{3}.
\end{aligned}
$$

☞ The calculation of variance:

$$
\begin{aligned}
E[U^4] &= \int_0^1 t^4\, dt = \frac{t^5}{5}\Big|_0^1 = \frac{1}{5} \\
Var(U^2) &= E[U^4] - \big(E[U^2]\big)^2 \\
&= \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.
\end{aligned}
$$

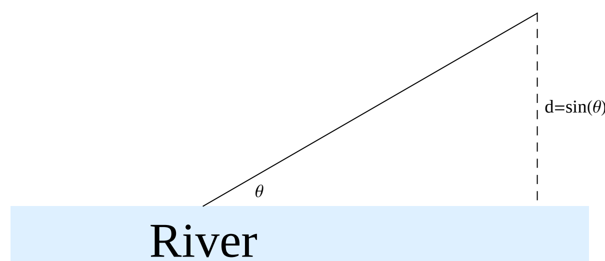




# Example

## Example

Suppose you walk the bank of a straight river and walk a kilometer in a randomly chosen direction. How far from the river are you likely to be.




Let $U$ be the uniform r.v. giving the distance angle from the shore, so the possible values are in the interval $[0, \pi]$.
Let $X$ be the distance from the shore, so $X = \sin(U)$.

## Example – continued

☞ The probability density for $U$ is $f_U(t) = \frac{1}{\pi}$ on $[0, \pi]$ and 0 elsewhere. So, the expected distance from the shore is

$$\begin{aligned} E[X] &= E[\sin(U)] = \int_0^\pi \sin(t)\frac{1}{\pi}\,dt \\ &= -\frac{1}{\pi}\cos(t)\Big|_0^\pi \\ &= \frac{2}{\pi} \approx 0.6366 \end{aligned}$$





## Example

**Example**

The Pilsdorff Beer Company runs a dilapidated fleet of trucks along a 100 mile road from Hangtown to Dry Gulch. The trucks are apt to break down at any point along the road with equal probability.

Where should the company locate the garage to minimize the distance they must tow the truck?

**Solution.** Let $b \in [0, 100]$ be the location of the garage, $U$ the uniform r.v. giving the location of the breakdown, and $X$ the distance traveled. So, $X = |U - b|$ and the density of $U$ is

$$f_U(t) = \begin{cases} \frac{1}{100} & \text{if } t \in [0, 100] \\ 0 & \text{otherwise.} \end{cases}$$

We want to minimize $E[X] = E[|U - b|]$.





## Example – continued

☞ The r.v. $X$ in terms of $U$ is

$$X = \begin{cases} b - U & \text{if } 0 \le U < b \\ U - b & \text{if } b \le U \le 100 \\ 0 & \text{otherwise.} \end{cases}$$

We compute expectation using the density of $f_U$:

$$\begin{aligned} E[X] &= \int_0^b (b-t)\frac{1}{100}\,dt + \int_b^{100} (t-b)\frac{1}{100}\,dt \\ &= \frac{1}{100}\Big((bt - \frac{t^2}{2})\big|_0^b + (\frac{t^2}{2} - bt)\big|_b^{100}\Big) \\ &= \frac{b^2 - \frac{b^2}{2}}{100} + 50 - b + \frac{b^2 - \frac{b^2}{2}}{100} \\ &= \frac{b^2}{100} - b + 50 \end{aligned}$$





## Example – continued

$$E[X] = \frac{b^2}{100} - b + 50$$

☞ We want to minimize $E[X]$. The expectation depends on $b$, and is a parabola opening upward. The minimum occurs when

$$0 = \frac{d}{db}E[X] = \frac{b}{50} - 1, \qquad \text{equivalently, } b = 50$$

☞ Pillsdorff should build their garage midway between Hangtown and Dry Gulch.