# Lecture 3: Random variables, distributions, and transformations

## Definition 1.4.1.

A random variable $X$ is a function from $S$ into a subset of $\mathscr{R}$ such that for any Borel set $B \subset \mathscr{R}$ $\{X \in B\} = \{\omega \in S : X(\omega) \in B\}$ is an event.

- It is convenient to deal with a summary variable than the elements in the original sample space.
- The range of a random variables is simpler than $S$.
- In some problems, there is a natural random variable; e.g., the number of accidents, number of successes, etc. In other cases, we may define a random variable according to our interests.

## Example 1.4.3. (Toss a coin 3 times)

| outcome | hhh | hht | hth | thh | tth | tht | htt | ttt |
|---|---|---|---|---|---|---|---|---|
| $X$: number of heads | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 0 |

- The range of $X$, $\{0, 1, 2, 3\}$, is simpler than $S$.
- $X$ treats hht, hth, thh the same, and tth, tht, htt the same.

## Induced probability

The induced probability of $X$ is

$$P_X(B) = P(X \in B) = P(\{\omega \in S : X(\omega) \in B\})$$

The probability $P_X$ is called the distribution of $X$.

## Example 1.4.3 (continued)

If the coin is fair, then

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

A table like this effective for $X$ taking finite many values.

## Definition 1.5.1.

The cumulative distribution function (cdf) of a random variable $X$, denoted by $F_X(x)$, is defined by

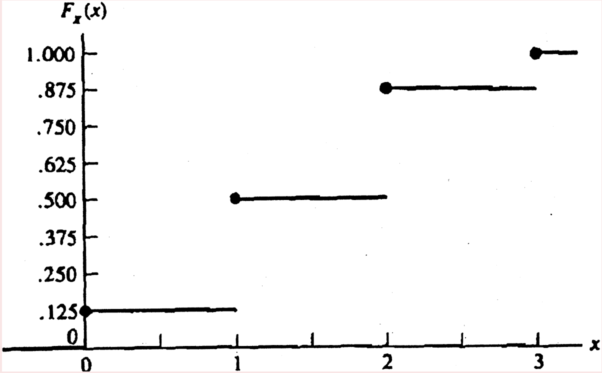
$$F_X(x) = P_X(X \leq x), \qquad x \in \mathscr{R}$$

Even if $X$ may not take all real numbers, the cdf is defined for all $x \in \mathscr{R}$

# Example 1.4.3 (continued)

The cdf of $X$ is a step function

$$F_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ 0.125 & 0 \le x < 1 \\ 0.5 & 1 \le x < 2 \\ 0.875 & 2 \le x < 3 \\ 1 & 3 \le x < \infty \end{cases}$$



## Theorem 1.5.3.

The function $F(x)$ is a cdf iff
a. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$;
b. $F(x)$ is nondecreasing in $x$;
c. $F(x)$ is right-continuous: $\lim_{\varepsilon\to 0, \varepsilon>0} F(x+\varepsilon) = F(x)$ for any $x \in \mathscr{R}$.

This theorem says that

- if $F$ is the cdf of a random variable $X$, then $F$ satisfies a-c (this is easy to prove);
- if $F$ satisfies a-c, then there exists a random variable $X$ such that the cdf of $X$ is $F$ (this is not easy to prove).

## Definition 1.5.7.

A random variable $X$ is continuous if $F_X(x)$ is continuous in $x$.
A random variable $X$ is discrete if $F_X(x)$ is a step function of $x$.

There are random variables that are neither discrete nor continuous; e.g., a mixture of the two types.

## Example

Let $c$ be a constant and

$$G_c(x) = \begin{cases} 0 & x < 0 \\ 1 - ce^{-3x} & x \geq 0 \end{cases}$$

For what values of $c$, $G_c$ is a cdf?

Clearly,

$$\lim_{x \to -\infty} G_c(x) = 0, \qquad \lim_{x \to \infty} G_c(x) = 1$$

$$G_c'(x) = \begin{cases} 0 & x < 0 \\ 3ce^{-3x} & x > 0 \end{cases}$$

- If $c < 0$, $G_c$ is decreasing.
- If $c > 1$, $G_c(0) = 1 - c < 0$.
- If $0 \leq c \leq 1$, then $G_c$ is nondecreasing.
- $G_c$ is always right-continuous.

Hence, $G_c$ is a cdf iff $0 \leq c \leq 1$.
If $c = 1$, then $G_c$ is continuous.
If $c = 0$, then $G_c$ is a special discrete cdf.
If $0 < c < 1$, then $G_c$ is neither continuous nor discrete.

## Definition 1.5.8/Theorem 1.5.10.

Two random variables $X$ and $Y$ with cdf's $F_X$ and $F_Y$ respectively are identically distributed iff $F_X(x) = F_Y(x)$ for any $x \in \mathscr{R}$, which is equivalent to $P(X \in B) = P(Y \in B)$ for any Borel set $B$.

Note that two identically distributed random variables are not necessarily equal (see Example 1.5.9).
The cdf's can be used to calculate various probabilities related to random variables but sometimes it is more convenient to use another function to calculate probabilities.

## Definition 1.6.1.

The probability mass function (pmf) of a discrete random variable $X$ is

$$f_X(x) = P(X = x) \qquad x \in \mathscr{R}$$

- For discrete $X$, $f_X(x) > 0$ for only countably many $x$'s.
- $f_X(x)$ is the size of the jump in the cdf $F_X$ at $x$.
- For any Borel set $A$,

$$P(X \in A) = \sum_{k \in A} f_X(k) \quad \text{and} \quad F_X(x) = P(X \leq x) = \sum_{k \leq x} f_X(k)$$

## How to define something similar to a pmf for a continuous $X$?

Defining $f_X(x) = P(X = x)$ does not work for a continuous $X$.
For any $\varepsilon > 0$,

$$P(X = x) \leq P(x - \varepsilon < X \leq x) = F_X(x) - F_X(x - \varepsilon)$$

Since $F_X$ is continuous at $x$, we obtain

$$P(X = x) \leq \lim_{\varepsilon \to 0} [F_X(x) - F_X(x - \varepsilon)] = 0$$

Hence, $P(X = x) = 0$ for any $x \in \mathscr{R}$ (compare this to a discrete $X$).
For a discrete $X$, we have

$$F_X(x) = \sum_{k \leq x} f_X(k)$$

An analog for the continuous case is to replace the sum by an integral.

### Definition 1.6.3.

The probability density function (pdf) of a continuous random variable $X$ (if it exists) is the function $f_X(x)$ satisfying

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \qquad x \in \mathscr{R}$$

- Unlike the discrete case, a pdf of a continuous $X$ may not exist.
- For a continuous $X$, it is convenient to use the pdf to calculate probabilities
- If $F_X$ is differentiable, then $f_X(x) = F'_X(x) = \frac{d}{dx} F_X(x)$.
- A continuous random variable has a pdf iff its cdf is absolutely continuous.
- If $f$ is a pdf, the set $\{x : f(x) > 0\}$ is called its support.

**Theorem 1.6.5.**

A function $f(x)$ is a pdf iff
a. $f(x) \geq 0$ for all $x$;
b. $\int_{-\infty}^{\infty} f(x)dx = 1$.

**Calculation of probabilities**

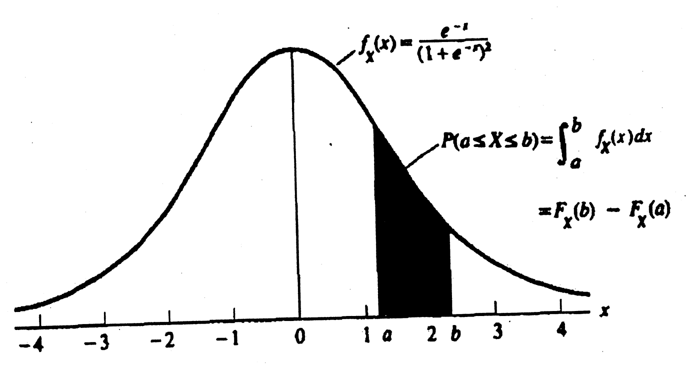
For a continuous $X$ with pdf $f_X$,

$$P(X \in A) = \int_A f_X(x)dx$$

For example,

$$P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)$$
$$= P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt$$

which is the area under the curve $f_X(t)$.



## Example (finding the cdf, given a pdf)

Suppose that

$$f_X(x) = \begin{cases} 1+x & -1 \leq x < 0 \\ 1-x & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$F_X(x) = ?$
Obviously, $F_X(x) = 0$ when $x < -1$.
For $-1 \leq x < 0$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-1}^{x} (1+t)dt = \frac{t^2}{2} + t\Big|_{t=-1}^{x} = \frac{x^2}{2} + x + \frac{1}{2}$$

For $0 \leq x < 1$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-1}^{0} (1+t)dt + \int_{0}^{x} (1-t)dt = \frac{1}{2} - \frac{x^2}{2} + x$$

For $x \geq 1$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-1}^{0} (1+t)dt + \int_{0}^{1} (1-t)dt = \frac{1}{2} + \frac{1}{2} = 1$$

## Finding the pdf, given a cdf

- If $F_X$ is differentiable everywhere, then $f_X = F'_X$.
- What if $F_X$ is differentiable except for some (at most countably many) points?
  $f_X(x) = F'_X(x)$ for $x$ at which $F_X$ is differentiable;
  $f_X(x)$ can be any $c \geq 0$ for $x$ at which $F_X$ is not differentiable.

## Example

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x/2 & 0 \leq x < 1/2 \\ x^2 & 1/2 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

$f_X(x) = ?$
$F_X$ is differentiable except at 0, 1/2 and 1.

$$f_X(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \leq x < 1/2 \\ 2x & 1/2 \leq x < 1 \\ 0 & x \geq 1 \end{cases}$$

## Example.

What values of $\theta$ and $\beta$ will make $f(x)$ a pdf?

$$f(x) = \beta e^{\theta x - |x|} \qquad x \in \mathscr{R}$$

If $\theta \geq 1$, then $\lim_{x \to \infty} e^{\theta x - |x|} = \infty$; if $\theta \leq -1$, then $\lim_{x \to -\infty} e^{\theta x - |x|} = \infty$; hence, $\theta$ has to be in $(-1, 1)$.

For $\theta \in (-1, 1)$,

$$\int_{-\infty}^{\infty} e^{\theta x - |x|} dx = \int_{-\infty}^{0} e^{\theta x + x} dx + \int_{0}^{\infty} e^{\theta x - x} dx = \frac{1}{1+\theta} + \frac{1}{1-\theta} = \frac{2}{1-\theta^2}$$

Hence, in orde to have a pdf, we must have $\beta = (1 - \theta^2)/2$.

What is the cdf of $f$?

For $x \leq 0$,

$$F(x) = \int_{-\infty}^{x} f(t) dt = \frac{1-\theta^2}{2} \int_{-\infty}^{x} e^{\theta t + t} dt = \frac{1-\theta}{2} e^{(1+\theta)x}$$

For $x > 0$,

$$F(x) = \int_{-\infty}^{x} f(t) dt = \frac{1-\theta^2}{2} \left[ \int_{-\infty}^{0} e^{\theta t + t} dt + \int_{0}^{x} e^{\theta t - t} dt \right] = 1 - \frac{1+\theta}{2} e^{(1-\theta)x}$$

## Transformations

For a random variable $X$, we are often interested in a transformation, $Y = g(X)$, which is also a random variable.
Here $g$ is a function from the space for $X$ to a new space.
For an event $A$ in the $Y$-space,

$$P(Y \in A) = P(g(X) \in A) = P(X \in g^{-1}(A))$$

where

$$g^{-1}(A) = \{x : g(x) \in A\}.$$

Here, $g^{-1}$ is not the inverse function and $g^{-1}(A)$ is an event.

## Example: $g(x) = x^2$

If $A = \{t\}$, a single point with $t \geq 0$, then $g^{-1}(\{t\}) = \{-\sqrt{t}, \sqrt{t}\}$, since $g(-\sqrt{t}) = (-\sqrt{t})^2 = t$ and $g(\sqrt{t}) = (\sqrt{t})^2 = t$.
If $A = [a, b]$ with $0 \leq a \leq b$, then $g^{-1}(A) = [-\sqrt{b}, -\sqrt{a}] \cup [\sqrt{a}, \sqrt{b}]$.

## How to obtain $F_Y$ (or $f_Y$) using $F_X$ (or $f_X$)?

If $X$ is discrete, then $Y$ is also discrete.
The pmf for $Y$ is

$$f_Y(y) = P(Y = y) = P(g(X) = y)$$

$$= \sum_{x:\ g(x)=y} P(X = x) = \sum_{x:\ x \in g^{-1}(\{y\})} f_X(x).$$

For discrete $X$, we often can directly work out a formula for $f_Y$.

### Example

Suppose $X$ has the following distribution

| $x$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|
| $f_X(x)$ | 0.1 | 0.2 | 0.1 | 0.2 | 0.3 | 0.1 |

Let $Y = X^2$.
Then $Y$ can take values 0, 1, 4, and 9 and the distribution of $Y$ is

| $y$ | 0 | 1 | 4 | 9 |
|---|---|---|---|---|
| $f_Y(y)$ | 0.1 | 0.4 | 0.4 | 0.1 |

## Continuous random variables

If $X$ is continuous with a pdf $f_X$, then

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{\{x:\ g(x) \leq y\}} f_X(x)dx$$

In general, the region $\{x : g(x) \leq y\}$ may be difficult to identify. Sometimes it is possible to find simple formulas for the cdf and pdf of $Y$ in terms of the cdf and pdf of $X$.

## Example: $Y = X^2$

If $y \leq 0$, $F_Y(y) = 0$ and $f_Y(y) = 0$.
If $y > 0$,

$$\begin{aligned} F_Y(y) &= \int_{\{x:\ x^2 \leq y\}} f_X(x)dx = \int_{\{x:\ -\sqrt{y} \leq x \leq \sqrt{y}\}} f_X(x)dx \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\ f_Y(y) &= \frac{d}{dy} F_Y(y) = f_X(\sqrt{y})\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\frac{1}{2\sqrt{y}} \end{aligned}$$

If $g$ is strictly monotone, then the inverse function $g^{-1}$ exists and

$$\{x : g(x) \le y\} = \begin{cases} \{x : x \le g^{-1}(y)\} & \text{if } g \text{ is increasing} \\ \{x : x \ge g^{-1}(y)\} & \text{if } g \text{ is decreasing} \end{cases}$$

## Theorem 2.1.3/Theorem 2.1.5

Let $X$ be a continuous random variable and $Y = g(X)$ with range $\mathscr{Y}$.

a. If $g$ is increasing, then $F_Y(y) = F_X(g^{-1}(y))$, $y \in \mathscr{Y}$.

b. If $g$ is decreasing, then $F_Y(y) = 1 - F_X(g^{-1}(y))$, $y \in \mathscr{Y}$.

c. If $f_X$ is continuous and $g$ is continuously differentiable, then

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| & y \in \mathscr{Y} \\ 0 & \text{otherwise} \end{cases}$$

## Example 2.1.5.

Suppose that $Y = g(X)$, $g(x) = 1/x$, and $X$ has the pdf

$$f_X(x) = [(n-1)!\beta^n]^{-1} x^{n-1} e^{-x/\beta}, \quad x > 0$$

Then $\mathscr{Y} = (0, \infty)$, $g^{-1}(y) = 1/y$, $\frac{d}{dy}g^{-1}(y) = -1/y^2$, and

$$f_Y(y) = f_X(g^{-1}(y))\left|-1/y^2\right| = [(n-1)!\beta^n]^{-1} y^{-(n+1)} e^{-1/(\beta y)}, \quad y > 0$$

Many useful functions are not monotone.
The next theorem extends Theorem 2.1.5 to the case where $g$ is piecewise monotone.

## Theorem 2.1.8.

Let $X$ be a continuous random variable with pdf $f_X$. Suppose that there are disjoint $A_1,...,A_k$ such that $P(X \in \cup_{t=0}^{k} A_t) = 1$, $f_X$ is continuous on each $A_t$, $t = 1,...,k$, and there are functions $g_1(x),...,g_k(x)$ defined on $A_1,...,A_k$, respectively, satisfying

i. $g(x) = g_t(x)$ for $x \in A_t$;
ii. $g_t(x)$ is strictly monotone on $A_t$;
iii. the set $\mathscr{Y} = \{y : y = g_t(x) \text{ for some } x \in A_t\}$ is the same for each $t$;
iv. $g_t^{-1}(y)$ has a continuous derivative on $\mathscr{Y}$ for each $t$.

Then $Y$ has the pdf

$$f_Y(y) = \begin{cases} \sum_{t=1}^{k} f_X(g_t^{-1}(y)) \left| \frac{d}{dy} g_t^{-1}(y) \right| & y \in \mathscr{Y} \\ 0 & \text{otherwise} \end{cases}$$

## Example 2.1.9 ($Y = X^2$)

Let $A_1 = (-\infty, 0)$, $A_2 = (0, \infty)$, $\mathscr{Y} = (0, \infty)$.
On $A_1$, $g_1(x) = x^2$ is decreasing, $g_1^{-1}(y) = -\sqrt{y}$;
On $A_2$, $g_2(x) = x^2$ is increasing, $g_2^{-1}(y) = \sqrt{y}$.
By Theorem 2.1.8,

$$f_Y(y) = f_X(-\sqrt{y}) \left| -\frac{1}{2\sqrt{y}} \right| + f_X(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| \quad y > 0$$

We obtained this previously.

As a special case, consider

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathscr{R} \quad \text{(the standard normal)}$$

Then, for $y > 0$,

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left| \frac{1}{2\sqrt{y}} \right| = \frac{1}{\sqrt{2\pi y}} e^{-y/2},$$

which is the chi-square pdf with 1 degree of freedom.