# Chapter 3 sections

- 3.1 Random Variables and Discrete Distributions
- 3.2 Continuous Distributions
- 3.3 The Cumulative Distribution Function
- 3.4 Bivariate Distributions
- 3.5 Marginal Distributions
- 3.6 Conditional Distributions
- 3.7 Multivariate Distributions (generalization of bivariate)
- 3.8 Functions of a Random Variable
- 3.9 Functions of Two or More Random Variables
- SKIP: 3.10 Markov Chains

# Random Variables

## Def: Random Variable

A *random variable* is a function from a sample space $S$ to the real numbers $\mathbb{R}$

$$P(X = x_i) = P(\{s_j \in S : X(s_j) = x_i\})$$

or

$$P(X \in A) = P(\{s_j \in S : X(s_j) \in A\})$$

- Note: Random variables are denoted by capital letters and the values they take (their outcome) with lowercase

Examples:

| Experiment | Random variable |
|---|---|
| Toss two dice | $X =$ sum of the numbers |
| Apply different amounts of fertilizer to corn plants | $X =$ yield per acre |

# Discrete random variables

## Def: Probability (mass) function

A random variable $X$ is said to have a *discrete distribution* if $X$ can only take countable number of different values.
The *probability function (pf)* for $X$ is defined as

$$f(x) = P(X = x) \quad \text{defined for all } x \in \mathbb{R}$$

## Bernoulli distribution

Let $p$ be the probability of winning a bet and define $X = 0$ if we loose and $X = 1$ if we win. Then $X$ has the *Bernoulli distribution with parameter $p$*, often denoted $X \sim \mathrm{Bernoulli}(p)$, and the pf is

$$f(x) = \begin{cases} p & \text{if } x = 0 \\ 1 - p & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

# Discrete random variables

## Def: Probability (mass) function

A random variable $X$ is said to have a *discrete distribution* if $X$ can only take countable number of different values.
The *probability function (pf)* for $X$ is defined as

$$f(x) = P(X = x) \quad \text{defined for all } x \in \mathbb{R}$$

## Bernoulli distribution

Let $p$ be the probability of winning a bet and define $X = 0$ if we loose and $X = 1$ if we win. Then $X$ has the *Bernoulli distribution with parameter p*, often denoted $X \sim \mathrm{Bernoulli}(p)$, and the pf is

$$f(x) = \begin{cases} p & \text{if } x = 0 \\ 1 - p & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

# Example: Rolling sixes

- We are interested in the number of 6's we obtain in four rolls of a fair dice.
- Find the pf of this random variable
- Use this pf to calculate the probability of obtaining at least one 6 in four rolls.

# Binomial distribution

## Binomial distribution

$X =$ the number of "successes" in $n$ independent trials, where the probability of success is $p$. Then

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \ldots, n$$

$$\text{so the pf is } f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{if } x = 0, 1, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

$X$ is said to have the *Binomial distribution with parameters n and p*, often denoted $X \sim \mathrm{Binomial}(n, p)$

- A $\mathrm{Binomial}(n, p)$ random variable is a sequence of $n$ independent $\mathrm{Bernoulli}(p)$ trials
- I.e. $\mathrm{Bernoulli}(p) = \mathrm{Binomial}(1, p)$

# Continuous random variables

## Def: Probability density function

A random variable $X$ is said to have a *continuous distribution* if there exists a non-negative function $f$ defined on the real line such that

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$$

The function $f$ is called the *probability density function (pdf)*.
The closure of the set $\{x : f(x) > 0\}$ is called the *support* of $X$.

Examples

- If $X$ is *Uniformly distributed in* $[a, b]$, or $X \sim \mathrm{Uniform}(a, b)$, every interval in $[a, b]$ has probability proportional to it's length. The pdf is

$$f(x) = \frac{1}{b-a}, \; x \in [a, b], \; 0 \text{ otherwise} .$$

# Properties of pdf's and pf's

## Theorem

A function $f(x)$ is a pdf (or pf) of a random variable $X$ if and only if both of the following holds

1. $f(x) \geq 0$ for all $x$
2. For pf's:

$$\sum_{i=1}^{\infty} f(x_1) = 1$$

for pdf's:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

The coefficient (e.g. $\frac{1}{b-a}$) that ensures that 2 is satisfied is called the *normalizing constant*.

# Examples

1. Show that

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n, \quad f(x) = 0, \text{ otherwise}$$

   is a pf

2. Show that

$$f(x) = \frac{1}{(1+x)^2}, \quad \text{for } x > 0, \quad f(x) = 0, \text{ otherwise}$$

   is a pdf

# Cumulative Distribution Function

## Def: Cumulative distribution function

The *Cumulative distribution function (cdf)* of a random variable $X$ is

$$F(x) = P(X \leq x) \quad \text{for} \; -\infty < x\infty$$

Relationship between the cdf and p(d)f

- Continuous distributions:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

$$\text{and} \qquad f(x) = \frac{d}{dx}F(x)$$

- Discrete distributions:

$$F(x_i) = P(X \leq x_i) = \sum_{\{u: u \leq x_i\}} f(u)$$

# Example

Sketch the pf and cdf for $\mathrm{Binomial}(4, 1/6)$ (Tossing sixes example)

$$f(x) = P(X = x) = \binom{4}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{4-x} \qquad x = 0, 1, 2, 3, 4$$

| x | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| f(x) | 0.482 | 0.386 | 0.116 | 0.015 | 0.001 |
| F(x) | 0.482 | 0.868 | 0.984 | 0.999 | 1.000 |

# Properties of the cdf

**Theorem**

A function $F(x)$ is a cdf if and only if the following three conditions hold:

1. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$
2. F(x) is a nondecreasing function of $x$
3. F(x) is right-continuous; i.e. $\lim_{x\downarrow x_0} F(x) = F(x_0)$

Note that $\lim_{x\uparrow x_0} F(x)$ is not necessarily equal to $F(x_0)$

- In practice we use the pdf (or pf) much more than the cdf.
- However, the cdf has some additional theoretical properties (e.g. uniqueness) that the pdf does not have.

# Example

Consider the following cdf of a random variable $X$:

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{1}{9}x^2 & \text{for } 0 < x \leq 3 \\ 1 & \text{for } x > 3 \end{cases}$$

1. Verify that this is a cdf
2. Find and sketch the pdf of $X$

# Properties of the cdf

**Theorem**

- $P(X > x) = 1 - F(x)$ for all $x$
- $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$ for all $x_1 < x_2$
- For all $x$: $P(X = x) = F(x) - F(x^-)$ where $F(x^-) = \lim_{y \uparrow x} F(y)$

- For discrete distributions $f(x) = P(X = x)$ is equal to the jump the cdf $F$ takes at $x$
- For continuous distributions $P(X = x) = 0 \neq f(x)$!

# Identically distributed

### Def: Identically distributed

Random variables $X$ and $Y$ are *identically distributed (id)* if for every set $A$ we have $P(X \in A) = P(Y \in A)$

- NOTE: $X$ and $Y$ are NOT necessarily the same
- Example: Let $X$ and $Y$ be the number of head and tails, respectively, in $n$ tosses of a fair coin. They are not the same random variable, but they have the same distribution!

### Theorem

The following are equivalent:

1. Random variables $X$ and $Y$ are identically distributed
2. $F_X(x) = F_Y(x)$ for all $x$

# Quantile Function

**Def: Quantiles/Percentiles**

Let $X$ be a random variable with cdf $F(x)$ and let $p \in (0, 1)$.
We define the *quantile function* of $X$ as

$$F^{-1}(p) = \text{the smallest } x \text{ such that } F(x) \geq p$$

$F^{-1}(p)$ is called the *p quantile* of $x$ or the $100 \times p$ *percentile* pf $X$

- If $F(x)$ is continuous and one-to-one the quantile function is the inverse of $F(x)$. Then there is only one $x$ such that $F(x) = p$.
- Example: The quantile function for
$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{1}{9}x^2 & \text{for } 0 < x \leq 3 \\ 1 & \text{for } x > 3 \end{cases} \quad \text{is} \quad F^{-1}(p) = 3\sqrt{p}, \quad p \in (0, 1)$$
- The *median* is sometimes defined as the 0.5 quantile, but sometimes as the midpoint of the interval $[x_1, x_2)$ where $F(x) = 0.5$ for all $x \in [x_1, x_2)$.