

version: uploaded 06192019

# Random Variables and Their Distributions

## Week 10

---

Yunkyu Sohn

School of Political Science and Economics

Waseda University

2019 Spring Statistics II

# Review of Week 9

---

- ▶ Probability (Chapter 6.1.)
  - ▶ Definitions
  - ▶ Axioms
  - ▶ Permutations
  - ▶ Combinations
- ▶ Conditional Probability (Chapter 6.2.)
  - ▶ Conditional probability
  - ▶ Joint probability
  - ▶ Independence

# Example 1: Write Down Equivalence/Definition and Plug in Values

---

► Rolling a dice once

► A: a multiple of 2; B: a multiple of 3

►  $\Omega =$

► as a set:  $A =$

$B =$

► as a set:  $A^c =$

$B^c =$

►  $P(A) =$

$P(B) =$

►  $P(A^c) =$

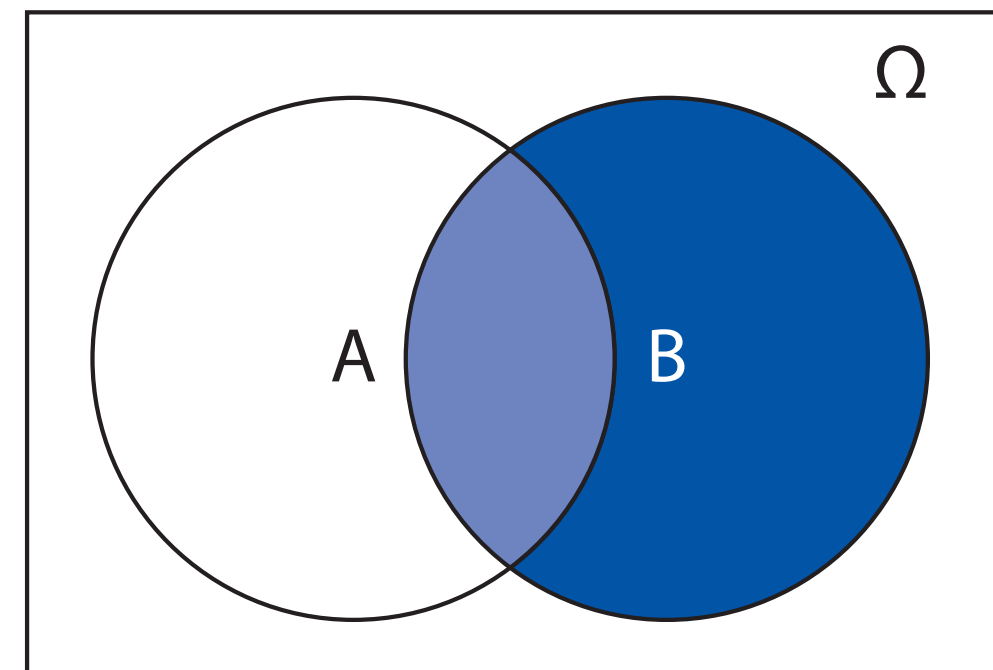
$P(B^c) =$

►  $P(A \text{ and } B) =$

►  $P(A|B) =$

►  $P(B|A) =$

►  $P(A) = P(A \text{ and } B) + P(A \text{ and } B^c)$



## Example 2: Write Down Equivalence/Definition and Plug in Values

- ▶ Flipping coin 3 times

- ▶ A: first flip head; B: second flip tail

- ▶  $\Omega =$

- ▶ as a set:  $A =$

- $B =$

- ▶ as a set:  $A^c =$

- $B^c =$

- ▶  $P(A) =$

- $P(B) =$

- ▶  $P(A^c) =$

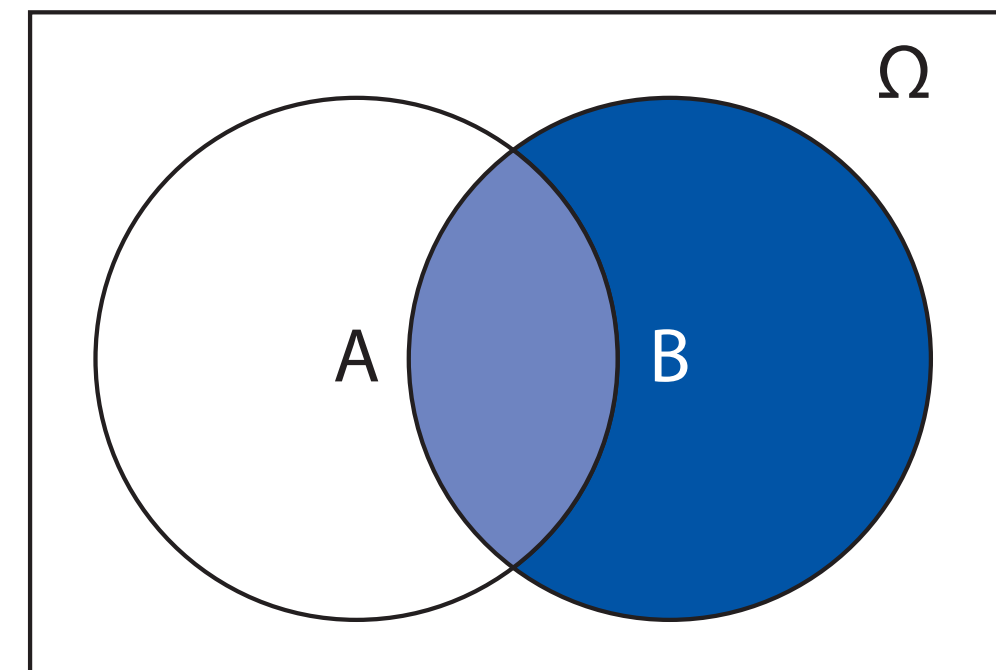
- $P(B^c) =$

- ▶  $P(A \text{ and } B) =$

- ▶  $P(A|B) =$

- ▶  $P(B|A) =$

- ▶  $P(A) = P(A \text{ and } B) + P(A \text{ and } B^c)$



# Example 3: Permutations and Combinations

---

- ▶ You are a producer in a large entertainment company
  - ▶ You want to form a 5-member unit from IZ\*ONE
    - ▶ Assume that you are selecting 5 completely at random
    - ▶ What is the chance that you select 1 JP member and 4 KR member?
    - ▶ What is the chance that you select at least 2 JP members?
- ▶ You are selecting 7 members from IZ\*ONE to cover a BTS song
  - ▶ Each will be assigned to a different role (e.g. V, Jin, RM ...)
    - ▶ How many potential scenarios are there?

# Bayes' Rule: Check Textbook 6.2.4

---

- Bayes' rule (obvious if you know definitions / axioms)

$$\begin{aligned} \underbrace{\Pr(A | B)}_{\text{conditional probability}} &= \frac{\overbrace{\Pr(A \text{ and } B)}^{\text{joint probability}}}{\underbrace{\Pr(B)}_{\text{marginal probability}}} \\ &= \frac{\Pr(B | A) \Pr(A)}{\Pr(B | A) \Pr(A) + \Pr(B | \text{not } A) \Pr(\text{not } A)} \end{aligned}$$

- What does  $\Pr(A|B)$  mean?
  - The chance of A being true given B is observed (i.e. true)
  - Why is above the most important result in probability theory?
    - e.g. **A is an event happened before B**; only B is observed
      - Using Bayes' rule: you can infer about A given B
- Bayesian updating: prior belief  $\Pr(A)$  -> posterior belief  $\Pr(A|B)$

# Bayes' Rule: Check Textbook 6.2.4

- Bayes' rule (obvious if you know definitions / axioms)

$$\underbrace{\Pr(A \mid B)}_{\text{conditional probability}} = \frac{\overbrace{\Pr(A \text{ and } B)}^{\text{joint probability}}}{\underbrace{\Pr(B)}_{\text{marginal probability}}} = \frac{\Pr(B \mid A) \Pr(A)}{\Pr(B \mid A) \Pr(A) + \Pr(B \mid \text{not } A) \Pr(\text{not } A)}$$

- **Posterior probability**  $\Pr(A|B)$ : belief of A after observing evidence B
- **Prior probability** (belief of A happening w/o evidence):  $P(A)$
- **Prior probability** (belief of A not happening w/o evidence):  $P(\text{not } A)$
- Examples (A: prior event; B: data/evidence)

A (prior event)	I voted for A	having a disease	studied hard
B (data)	A won	positive on medical test	high grade in exam

- Time order is not necessary but good for examples

# Example: Monty Hall Problem

---

- ▶ The most famous probability problem
  - ▶ <http://www.youtube.com/watch?v=mhlc7peGlGg>



- ▶ Even a great mathematician failed!



# Example: Monty Hall Problem

---

- ▶ You pick door A. Monty opens door C that has a goat.
  - ▶ Should you switch to door B?
- ▶ Prior beliefs:  $P(A) = P(B) = P(C) = 1/3$  (e.g. A: the car is behind door A)
- ▶ Data: Monty reveals door C (i.e. MC)
- ▶ Posterior belief (inferential goals):  $P(A \mid MC)$  and  $P(B \mid MC)$
- ▶ Question:  $P(A \mid MC) < P(B \mid MC)$
- ▶ What do we need? Bayes' rule!
  - ▶ Key: Monty's behavior is constrained to open a goat door
- ▶ Take home message: Do not believe your intuition, rely on logic.

# Contents

---

- ▶ Bayes' rule (Chapter 6.2.3)
- ▶ Random Variables and Probability Distributions (Chapter 6.3)
  - ▶ Overview
  - ▶ Bernoulli and uniform distributions
  - ▶ Binomial distribution
  - ▶ Uniform distribution
  - ▶ Normal (or Gaussian) distribution
  - ▶ Expectation

# Random Variables and Probability Distribution

---

- ▶ **Random variable** assigns a **numeric value** to each **event** of the experiment.
  - ▶ Coin flip side: head = 1; tail = 0
  - ▶ #secs took for commuting: any value greater than 0.
- ▶ These values represent **mutually exclusive and exhaustive events**, together forming the entire sample space  $\Omega$ .
- ▶ A **discrete RV** takes a finite or at most countably infinite #distinct values.
  - ▶ coin flip; race; number of years of education
- ▶ A **continuous RV** assumes an uncountably infinite number of values.
  - ▶ height, distance from earth, gross domestic product
- ▶ Probability distribution: Probability that a random variable takes a certain value or range of values.
  - ▶  $P(\text{side}): P(\text{side}=1) = 0.5; P(\text{side}=0) = 0.5$
  - ▶  $P(\text{\#secs}): P(0 < \text{\#secs} < 1000) = 0.3; P(1000 < \text{\#secs} < 2000) = 0.4 \dots$

# Probability Density / Mass Function

---

- ▶ Probability density function (PDF):  $f(x)$  for a continuous random variable
- ▶ Probability mass function (PMF):  $f(x)$  for a discrete random variable
- ▶ Cumulative density function (CDF):  $F(x)=P(X\leq x)$
- ▶ Cumulative mass function (CMF):  $F(x) = P(X \leq x) = \sum_{K \leq x} f(k)$ 
  - ▶ What is the probability that a random variable  $X$  takes a value equal to or less than  $x$ ?
  - ▶ Area under the density curve
  - ▶ Non-decreasing

# Bernoulli Distribution

---

►  $\Omega = \{0, 1\}$

► PMF

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

► CMF

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - p & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$



# Bernoulli Distribution

►  $\Omega = \{0, 1\}$

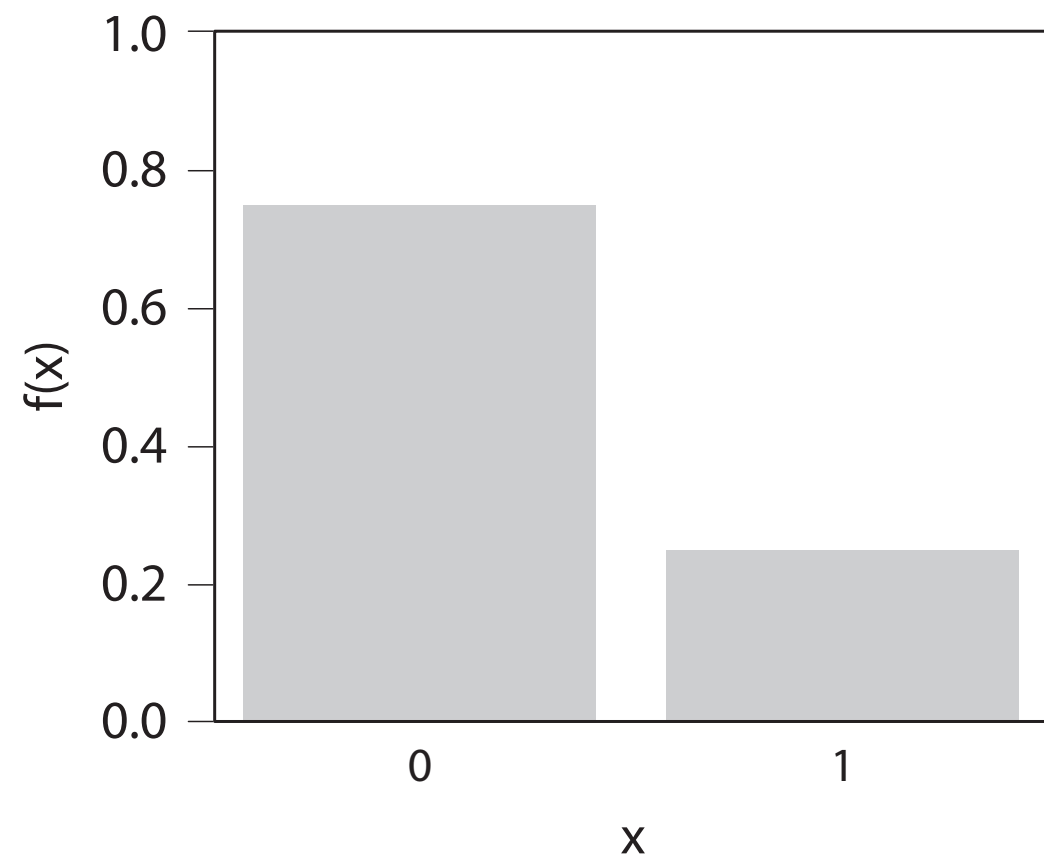
► PMF

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

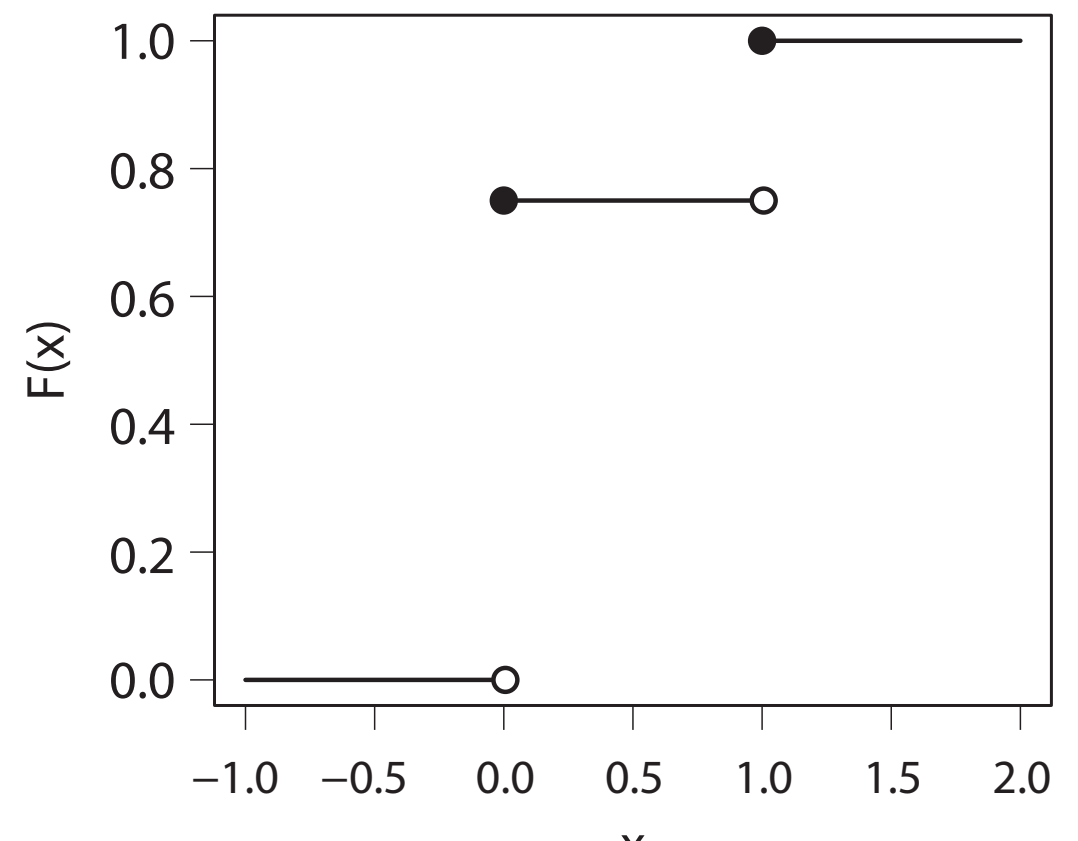


► PMF and CDF of Bernoulli distribution for  $p = 0.25$

Probability mass function

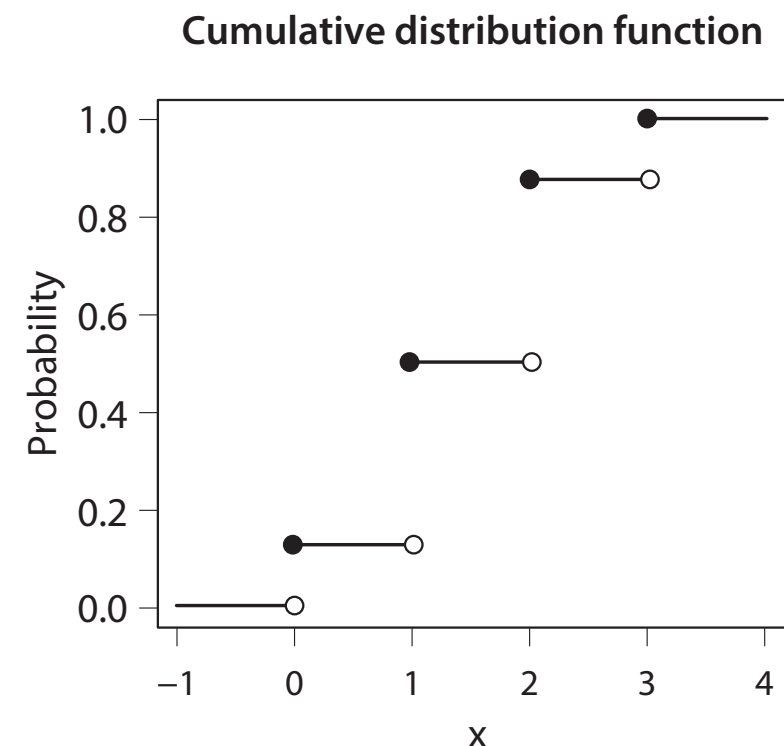
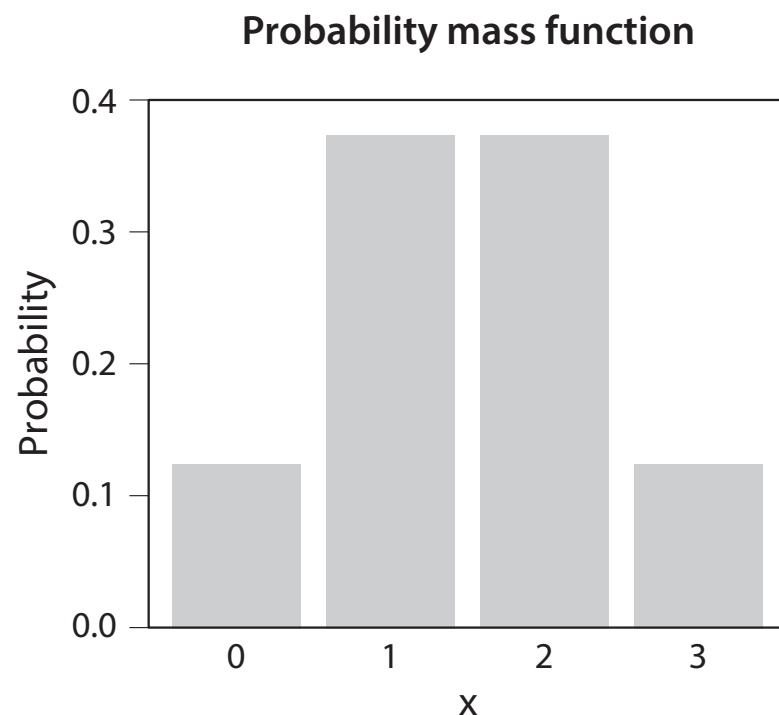


Cumulative distribution function



# Binomial Distribution

- ▶ The number of 1s (one of the binary outcomes) in **multiple** Bernoulli trials
- ▶  $\Omega = \{0, 1, \dots, n-1, n\}$
- ▶ PMF
$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \binom{n}{x} = {}_n C_x$$
- ▶ CMF
$$F(x) = P(X \leq x) = \sum_{k=0}^x \binom{n}{k} p^k (1 - p)^{n-k}$$
- ▶  $p = 0.5$  and  $n = 3$



# Binomial Distribution

---

- ▶ ex1) In a small department: There are exactly 10 who support candidate A, another 10 people who support candidate B for electing the chair of the department. Suppose that we expect their individual turnout probability is equal to their previous overall turnout rate which was 70%. What is the chance that exactly 7 people vote for candidate A and 7 people vote for candidate B, and the election ends in a tie?
- ▶ ex2) In a small department: There are exactly 10 who support candidate A, another 10 people who support candidate B for electing the chair of the department. Suppose that we expect their individual turnout probability is equal to their previous overall turnout rate which was 70%. What is the chance that the election ends in a tie?



# Uniform Distribution

- ▶ Every number in an interval has an equal chance of appearance

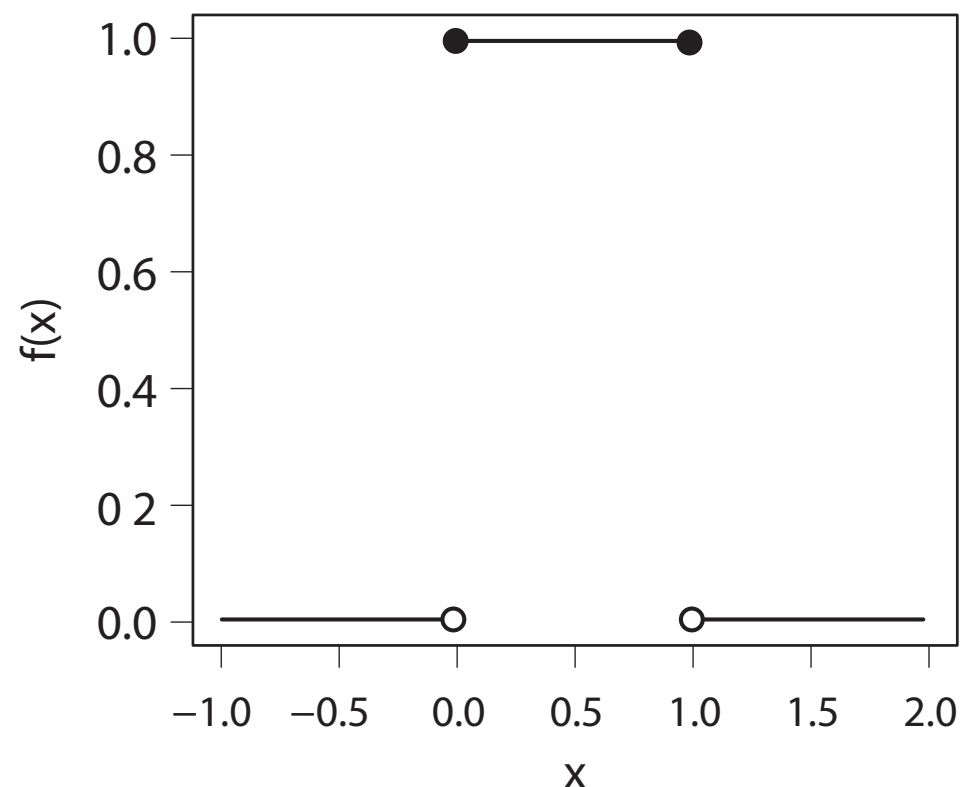
- ▶  $\Omega = \{x \mid a \leq x \leq b\}$

- ▶ PDF
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

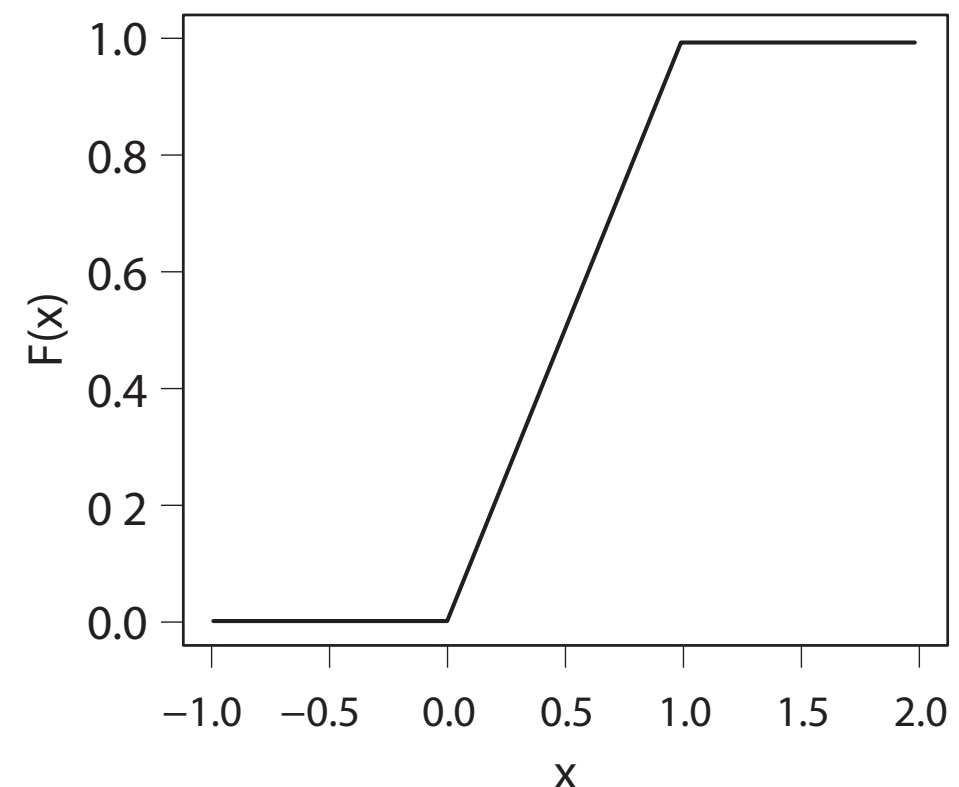
- ▶ CDF
$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & x \geq b \end{cases}$$

- ▶ Uniform distribution for the interval  $[0,1]$

Probability density function



Cumulative distribution function

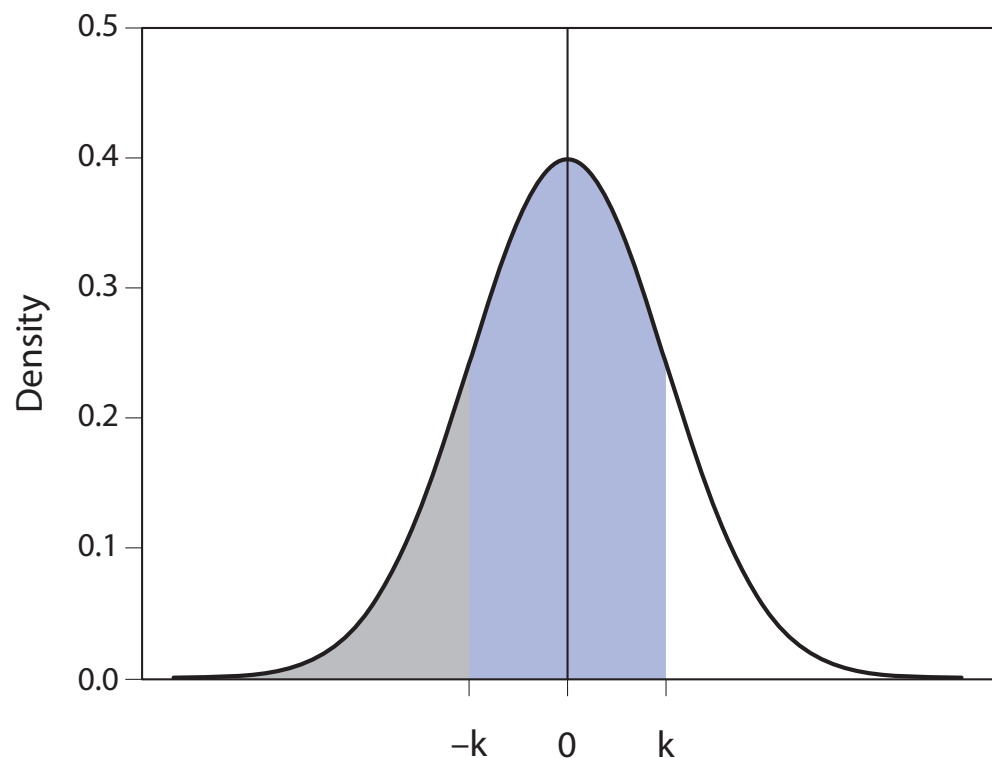


# Normal Distribution

- ▶ Most famous and frequently observed distribution (Why? -> next week)
- ▶  $\Omega$ =real numbers (continuous number)
- ▶  $X$  is normal RV with mean  $\mu$  and standard deviation  $\sigma$ :  $X \sim \mathcal{N}(\mu, \sigma^2)$

- ▶ PDF 
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}$$

- ▶  $X \sim \mathcal{N}(0,1)$



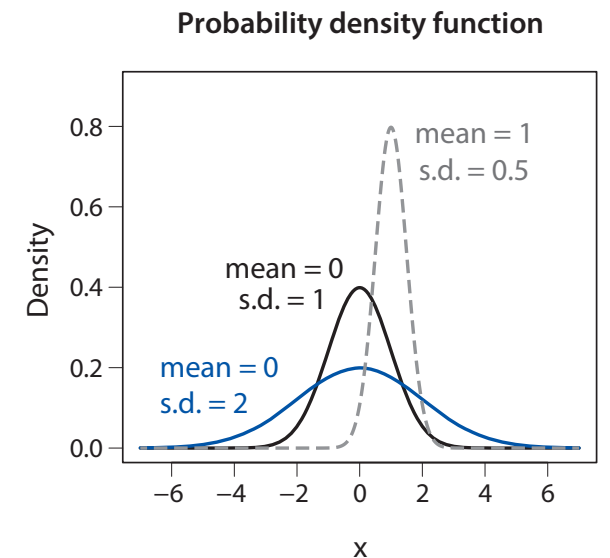
- ▶ Singled peaked, symmetric
- ▶ about 2/3 are within 1 standard deviation ( $\sigma$ ) from the mean
- ▶ about 95% are within 2 standard deviations ( $2\sigma$ ) from the mean

- ▶ CDF 
$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (t - \mu)^2 \right\} dt$$

# Normal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}$$

- ▶  $X \sim \mathcal{N}(\mu, \sigma^2)$
- ▶  $S = X + c \rightarrow S \sim \mathcal{N}(\mu + c, \sigma^2)$
- ▶  $Y = cX \rightarrow Y \sim \mathcal{N}(c\mu, (c\sigma)^2)$
- ▶ **z-score:**  $Z = (X - \mu)/\sigma \rightarrow Z \sim \mathcal{N}(0,1)$



- ▶ Probability that a normal random variable with mean  $\mu$  and sdv  $\sigma$  lies within  $k$  standard deviations from the mean for a positive constant  $k > 0$

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) = P(-k\sigma \leq X - \mu \leq k\sigma)$$

$$= P\left(-k \leq \frac{X - \mu}{\sigma} \leq k\right)$$

$$= P(-k \leq Z \leq k),$$

$$P(-k \leq Z \leq k) = P(Z \leq k) - P(Z \leq -k) = F(k) - F(-k)$$

# Normal Distribution

---

- ▶ Singled peaked, symmetric
- ▶ about 2/3 are within 1 standard deviation ( $\sigma$ ) from the mean
- ▶ about 95% are within 2 standard deviations ( $2\sigma$ ) from the mean

```
## plus minus 1 standard deviation from the mean
```

```
pnorm(1) - pnorm(-1)
```

```
## [1] 0.6826895
```

```
## plus minus 2 standard deviations from the mean
```

```
pnorm(2) - pnorm(-2)
```

```
## [1] 0.9544997
```

```
mu <- 5
```

```
sigma <- 2
```

```
## plus minus 1 standard deviation from the mean
```

```
pnorm(mu + sigma, mean = mu, sd = sigma) - pnorm(mu - sigma, mean = mu, sd = sigma)
```

```
## [1] 0.6826895
```

```
## plus minus 2 standard deviations from the mean
```

```
pnorm(mu + 2*sigma, mean = mu, sd = sigma) - pnorm(mu - 2*sigma, mean = mu, sd = sigma)
```

```
## [1] 0.9544997
```

# Expectation: Definition and General Properties

---

- Expectation of a random variable  $X$

$$\mathbb{E}(X) = \begin{cases} \sum_x x \times f(x) & \text{if } X \text{ is discrete,} \\ \int x \times f(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

- Properties of expectation ( $a, b$ : constant–fixed value–)

1.  $\mathbb{E}(a) = a$ .
2.  $\mathbb{E}(aX) = a\mathbb{E}(X)$ .
3.  $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$ .
4.  $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ .
5. If  $X$  and  $Y$  are independent, then  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ . But generally,  $\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y)$ .

# Expectation

---

- Expectation (mean) revisited

- Expected value of a random variable

- PMF of any discrete random variable

$$\mathbb{E}(X) = 0 \times f(0) + 1 \times f(1) + \cdots + n \times f(n) = \sum_{x=0}^n x \times f(x)$$

- PMF e.g. Bernoulli random variable

$$\mathbb{E}(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = 0 \times f(0) + 1 \times f(1) = 0 \times (1 - p) + 1 \times p = p$$

- PMF e.g. Binomial random variable (Y is a Bernoulli RV with  $p$ )

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n \mathbb{E}(Y_i) = np$$

- PDF for a continuous random variable defined in the interval  $[a, b]$

$$\mathbb{E}(X) = \int_a^b x \times f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{a+b}{2}$$

# Summary

---

- ▶ Bayes' rule (Chapter 6.2.3)
- ▶ Random Variables and Probability Distributions (Chapter 6.3)
  - ▶ Overview
  - ▶ Bernoulli and uniform distributions
  - ▶ Binomial distribution
  - ▶ Uniform distribution
  - ▶ Normal (or Gaussian) distribution
  - ▶ Expectation

**See you next week.**