A brief review of basics of probabilities

Milos Hauskrecht
milos@pitt.edu
5329 Sennott Square

Probability theory

Studies and describes random processes and their outcomes

- **Random processes may result in multiple different outcomes**
- **Example 1: coin flip**
  - Outcome is either head or tail (binary outcome)
  - Fair coin: outcomes are equally likely

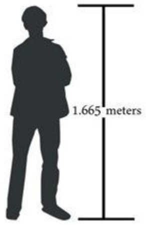
- **Example 2: sum of numbers obtained by rolling 2 dice**
  - Outcome number in between 2 to 12
  - Fair dices: outcome 2 is less likely then 3

Probability theory

Studies and describes random processes and their outcomes

- **Random processes may have multiple different outcomes**

- **Example 3: height of a person**
  - Select randomly a person from your school/city and report her height
  - Outcomes can be real numbers



- **And many others related to measurements, lotteries, etc**

Probabilities

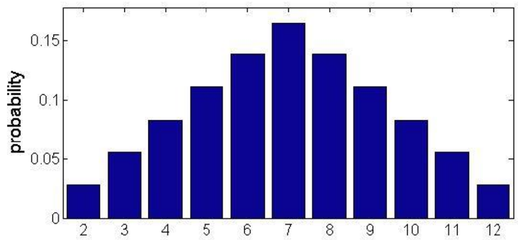
When the process is repeated many times outcomes occur with certain relative frequencies or **probabilities**

- **Example 1: coin flip**
  - **Fair coin:** outcomes are equally likely
    - Probability of head is 0.5 and tail is 0.5
  - Biased coin
    - Probability of head is 0.8 and tail is 0.2
    - Head outcome is 4 times more likely than tail

Probabilities

When the process is repeated many times outcomes occur with certain relative frequencies or **probabilities**

- **Example 2: sum of numbers obtained by rolling 2 dice**
  - Outcome number in between 2 to 12
  - Fair dice: outcome 2 is less likely then 3
    4 is less likely then 3, etc



Probability distribution function

Discrete (mutually exclusive) outcomes – the chance of outcomes is represented by **a probability distribution function**

- **probability distribution function – assigns a number between 0 and 1 to every outcome**
- **Example 1: coin flip**
  - Biased coin
    - Probability of head is 0.8 and tail is 0.2
    - Head outcome is 4 time more likely than tail

P(tail) = 0.2
P(head) = 0.8

$$P(coin) = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

- **What is the condition we need to satisfy ?**

Probability distribution function

Discrete (mutually exclusive) outcomes – the chance of outcomes is represented by **a probability distribution function**

- **probability distribution function – assigns a number between 0 and 1 to every outcome**
- **Example 1: coin flip**
  - Biased coin
    - Probability of head is 0.8 and tail is 0.2
    - Head outcome is 4 time more likely than tail

P(tail) = 0.2
P(head) = 0.8

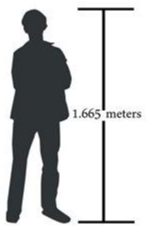
$$P(coin) = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

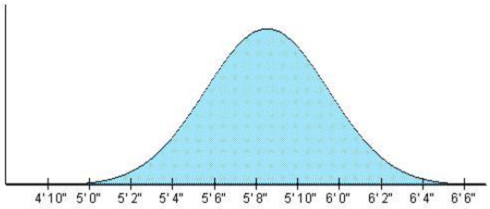
- **What is the condition we need to satisfy ?**
- **Sum of probabilities for discrete set of outcomes is 1**

Probability for real-valued outcomes

When the process is repeated many times outcomes occur with certain relative frequencies or **probabilities**

- **Example 3: height of a person**
  - Select randomly a person from your school/city and report her height
  - Outcomes can be real numbers
  - Different outcomes can be more or less likely



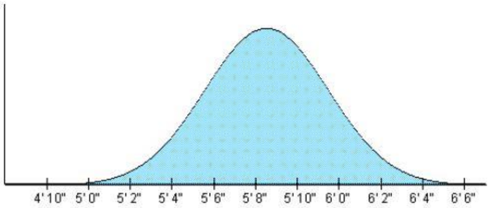


Normal (Gaussian) density

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**

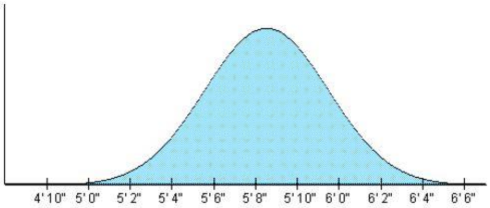


- **Condition on p(x) and 1?**

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**



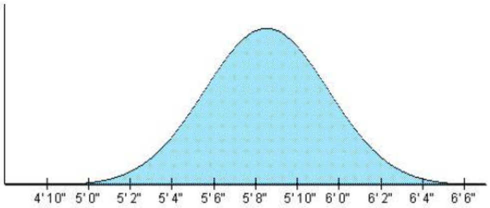
- **Condition on p(x) and 1?**

$$\int p(x)dx = 1$$

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**

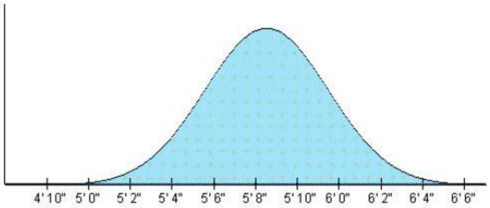


- **Can p(x) values for some x be negative?**

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**

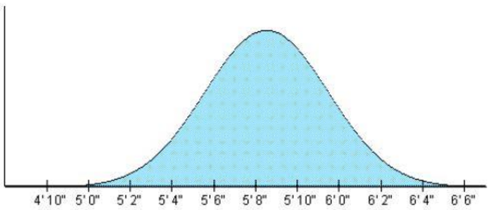


- **Can p(x) values for some x be negative?**
- **No**

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**

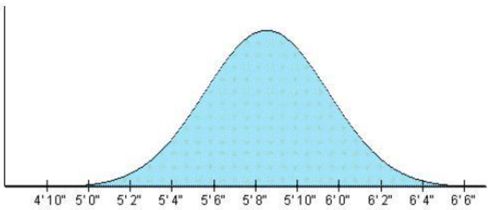


- **Can p(x) values for some x be > 1?**
  - **Remember we need** $\int p(x)dx = 1$

Probability density function

Real-valued outcomes – the chance of outcomes is represented by **a probability density function**

- **probability density function – p(x)**



- **Can p(x) values for some x be > 1?**
- **Remember we need:** $\int p(x)dx = 1$
- **Yes**

Random variable

Random variable = A function that maps observed outcomes (quantities) to real valued outcomes

Binary random variables: **Two outcomes** mapped to **0,1**

Example: Coin flip. Tail mapped to 0, Head mapped to 1

Note: Only one value for each outcome: either 0 or 1

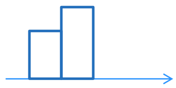
probability of tail $P(x=0)$

probability of head $P(x=1)$

Probability distribution: Assigns a probability to each possible outcome

A Biased coin

$P(x) = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}$

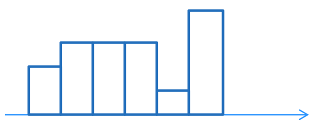


Random variable

Example: roll of a dice

- Outcomes =1,2,3,4,5,6 based on the roll of a die
- **trivial map to the same number**

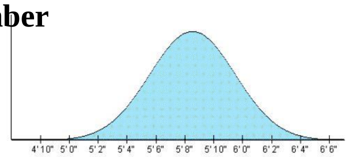
Biased dice



Example: x height of a person

Real valued outcomes

– trivial map to the same number



Probability

- Let A be an outcome event, and ¬A its complement.
  - **Then**

$$P(A) + P(\neg A) = ?$$

Probability

- Let A be an event, and ¬A its complement.
  - **Then**

$$P(A) + P(\neg A) = 1$$

$$P(A \wedge \neg A) = ?$$

Probability

- Let A be an event, and ¬A its complement.
  - Then

$$P(A) + P(\neg A) = 1$$

$$P(A \wedge \neg A) = 0$$

$$P(False) = 0$$

$$P(A \vee \neg A) = ?$$

Probability

- Let A be an event, and ¬A its complement.
  - Then

$$P(A) + P(\neg A) = 1$$

$$P(A \wedge \neg A) = 0$$

$$P(False) = 0$$

$$P(A \vee \neg A) = 1$$

$$P(True) = 1$$

Joint probability

Joint probability:

- **Let A and B be two events.** The probability of an event A, B occurring jointly

$$P(A \wedge B) = P(A, B)$$

We can add more events, say, A,B,C

$$P(A \wedge B \wedge C) = P(A, B, C)$$

Independence

Independence :

- Let A, B be two events. The events are independent if:

$$P(A, B) = ?$$

Independence

Independence :

- Let A, B be two events. The events are independent if:

$$P(A,B) = P(A)P(B)$$

Conditional probability

Conditional probability :

- Let A, B be two events. The conditional probability of A given B is defined as:

$$P(A \mid B) = ?$$

Conditional probability

Conditional probability :

- Let A, B be two events. The conditional probability of A given B is defined as:

$$P(A \mid B) = \frac{P(A,B)}{P(B)}$$

Product rule:

- A rewrite of the conditional probability

$$P(A,B) = P(A \mid B)P(B)$$

Bayes theorem

Bayes theorem

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Why?

$$P(A \mid B) = \frac{P(A,B)}{P(B)} \qquad P(A,B) = P(B \mid A)P(A)$$

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$