Discrete Random Variables

Jakob Bernoulli

Some random experiment outcomes are naturally associate with discrete outcomes such as number of boys in a family or number of defect light bulbs in a production series. Since it can not be predicted with certainty its called random variable.

Random variables

- A random variable is a numerical quantity that is generated by a random experiment.
- It will be denoted by capital letters, such as X or Z, and the actual values that they can take by lowercase letters, such as x and z.
- Examples for discrete random variables:

Experiment	Number *X*	Possible Values of *X*
Roll two fair dice	Sum of the number of dots on the top faces	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Flip a fair coin repeatedly	Number of tosses until the coin lands heads	1, 2, 3,4, ...

Discrete and continuous random variables

- A random variable is called **discrete** if it has either a finite or a countable number of possible values. They typically arise from counting.
- A random variable is called **continuous** if its possible values contain a whole interval of numbers. These typically arise from a measurement. Examples:

Experiment	Number *X*	Possible Values of *X*
Measure the voltage at an electrical outlet	Voltage measured	$118 \leq x \leq 122$
Operate a light bulb until it burns out	Time until the bulb burns out	$0 \leq x < \infty$

Probability distributions for discrete random variables

- Each possible value x of a discrete random variable X is the probability P(x) that X will take the value x in one trial of the experiment.
- The probability distribution of a discrete random variable X is a list of each possible value of X and their probability.
- The probabilities in the probability distribution of a random variable X must satisfy:
  - 1. Each probability P(x) must be between 0 and 1: $0 \leq P(x) \leq 1$.
  - 2. The sum of all the probabilities is 1: $\Sigma P(x) = 1$.

Example 1

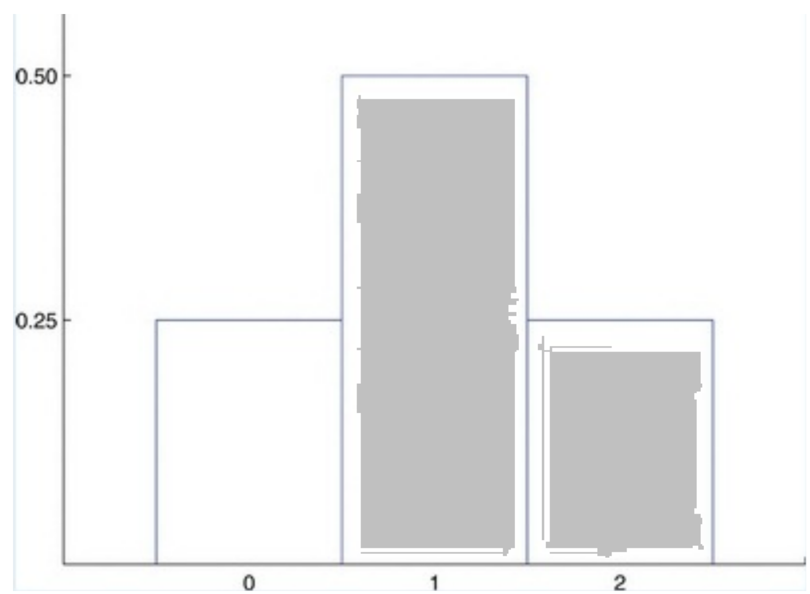
- We toss a coin twice. Let X be the number of heads observed.
- Probability distribution of X:

  S={hh,ht,th,tt} from counting:

x	0	1	2
$P(x)$	0.25	0.50	0.25

- Probability of observing at least one head?

  $P(X \geq 1)=P(1)+P(2)=0.75$



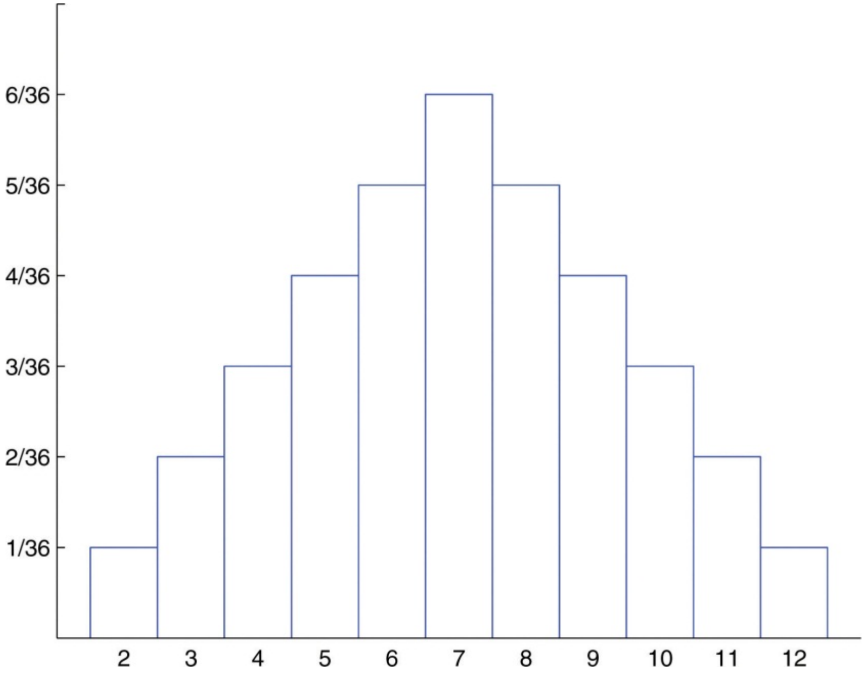
Probability distributions for discrete random variables

- Each possible value x of a discrete random variable X is the probability P(x) that X will take the value x in one trial of the experiment.
- The probability distribution of a discrete random variable X is a list of each possible value of X and their probability.
- The probabilities in the probability distribution of a random variable X must satisfy:
  - 1. Each probability P(x) must be between 0 and 1: $0 \leq P(x) \leq 1$.
  - 2. The sum of all the probabilities is 1: $\Sigma P(x) = 1$.

Example 2

- Two fair dice are rolled. Let X be the sum of both.
- Probability distribution:
- What it the probability

  for $P(X \geq 9)$?
- P(X is even)



The mean of a discrete random variable

- The mean (also called the expected value) of a discrete random variable X is the number: $\mu=E[X]=\Sigma x\ P(x)$
- The mean of a random variable may be interpreted as the average of the values assumed by the random variable in repeated trials of the experiment.
- Let X be the result of fair die, what it the expected value ?
- $E[X]= 1\frac{1}{6}+2\frac{1}{6}+3\frac{1}{6}+4\frac{1}{6}+5\frac{1}{6}+6\frac{1}{6} = (1+2+3+4+5+6)/6= 3.5$

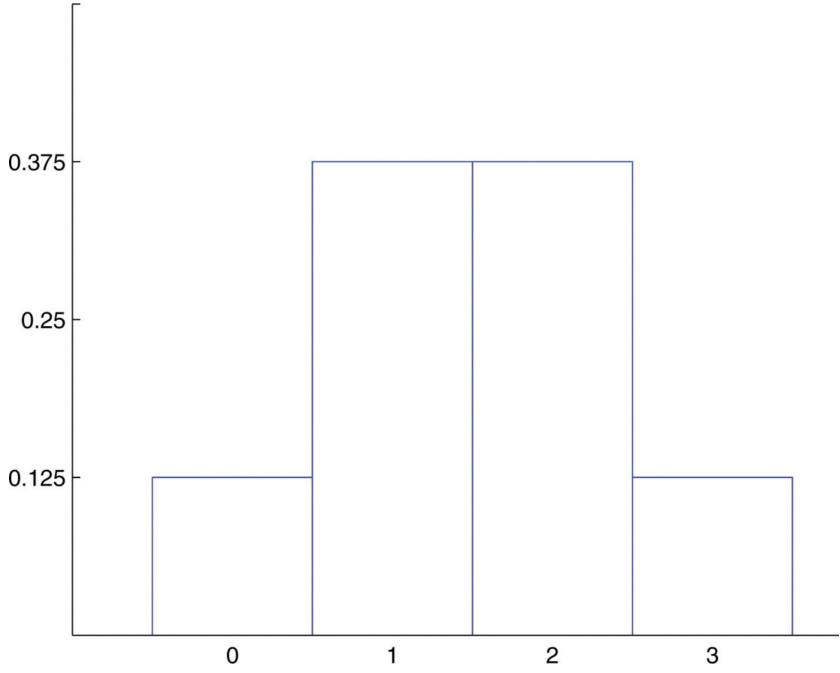
The standard deviation of a discrete random variable

- The variance σ^2 of a discrete random variable X is: $\sigma^2 = \Sigma(x-\mu)^2 P(x) = E[(X-E[X])^2]$

- The standard deviation, σ, of a discrete random variable X is the square root of its variance.

The binomial distribution

- Counting the number of successes such as for a series of three fair coin tosses or a gender of children in families with three children
- All three trials of the experiment have to be identical and independent, where each trial has two outcomes. The random variable that is generated is called the binomial random variable with parameters n=3 for the repetitions of the experiment and p=0.5 for the success probability.



The binomial distribution

- Suppose a random experiment has the following characteristics.

  1. There are n identical and independent trials of a common procedure.

  2. There are exactly two possible outcomes for each trial, one termed "success" and the other "failure."

  3. The probability of success on any one trial is the same number p.

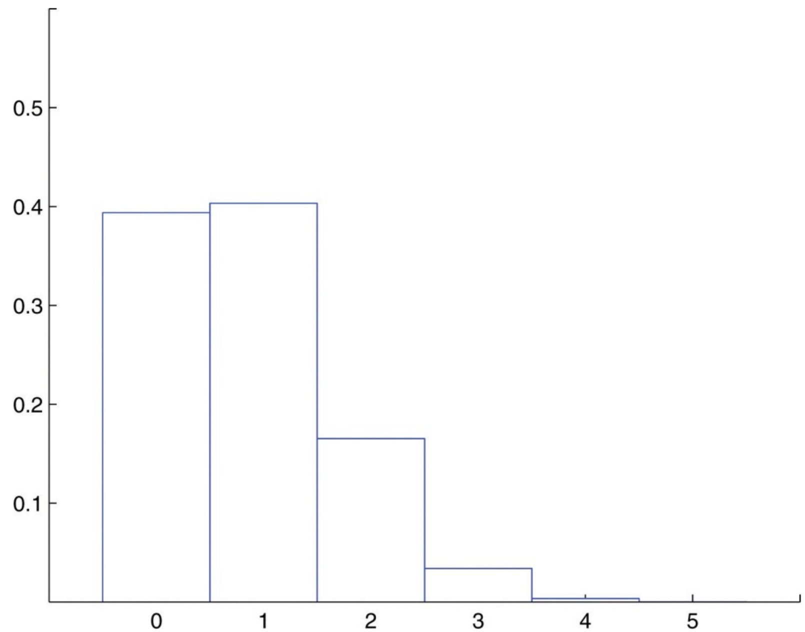
- Then the discrete random variable X that counts the number of successes in the n trials is the binomial random variable with parameters n and p. We also say that X has a binomial distribution with parameters n and p .
- Pdf of the binomial distribution: P(k)= $\binom{n}{k} p^k (1-p)^{n-k}$

  where k is the number of successes.

Example

- 17 percent of all fraud victims know the perpetrator of the fraud personally.
- Construct the probability distribution for a sample of 5.

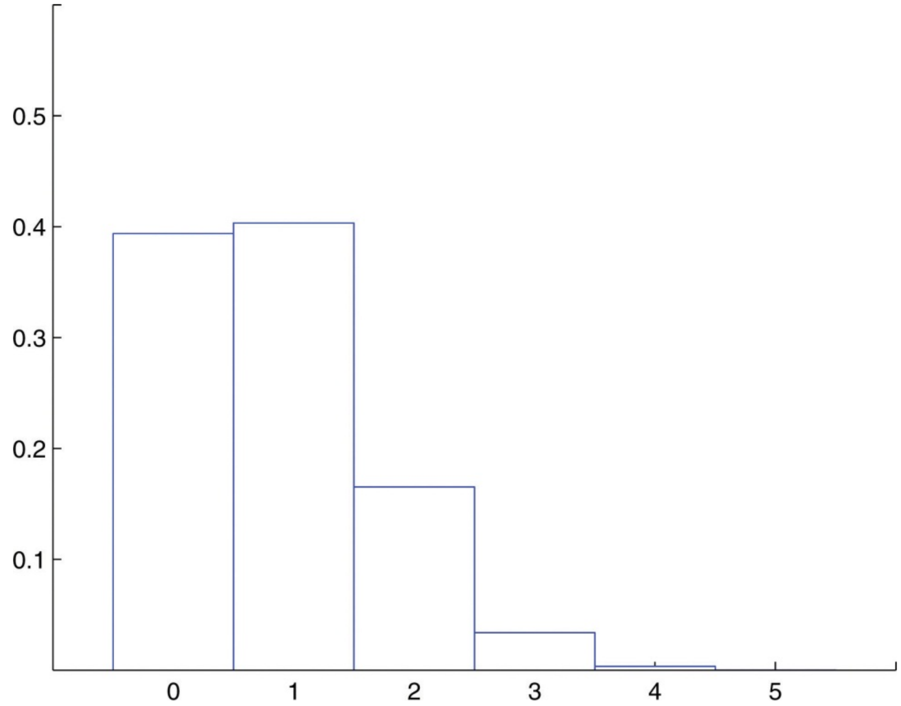
  - Binomial distributed with n=5,p=0.17



Example continued

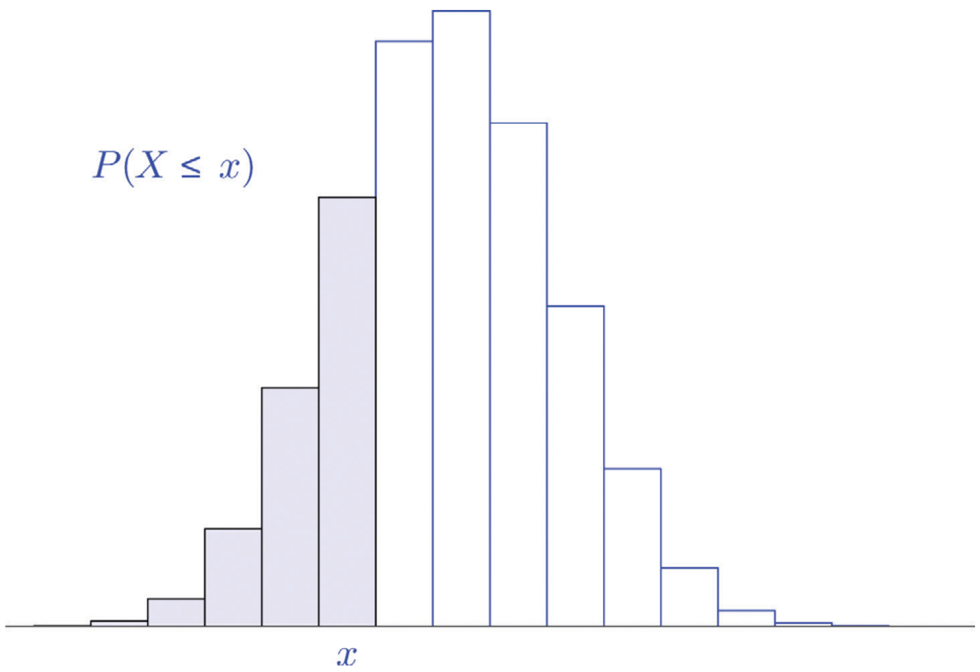
- An investigator examines 5 cases a day. What is the most frequent number of cases where the victim knew the perpetrator.

  -The number that is most likely is 1.



The cumulative probability distribution

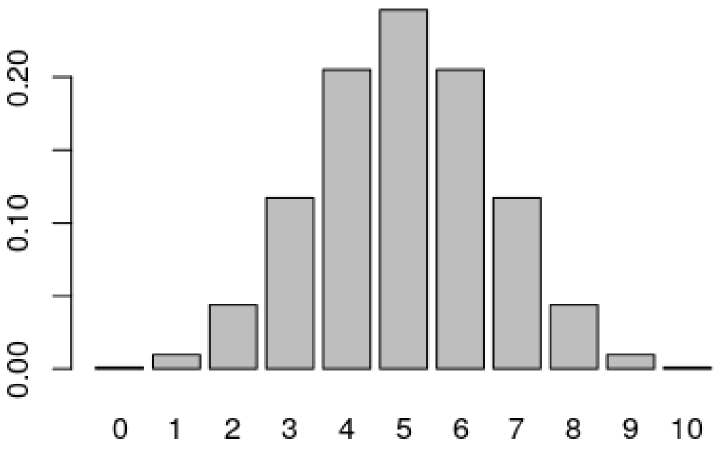
- The cumulative distribution function (CDF) of a random variable X, evaluated at x, is the probability that X will take a value less than or equal to x.
- $P(X \leq x) = P(0) + P(1) + \cdots + P(x)$



R exercise

- 10 coin tosses. Let X be number of heads occurred. Binomial distributed n=10, p=0.5

probability distribution of X



CDF of X