

L17: Random Variables

- Reading: Rosen 7.4

Random Variables

- **Definition**

A **random variable** is a function from the sample space of an experiment to the set of real numbers, i.e., a random variable assigns a real number to each possible outcome.

- Note: This function is not random!

- **Example**

Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. Then $X(t)$ is the function where:

$$X(HHH) = 3, X(TTT) = 0,$$

$$X(HHT) = X(HTH) = X(THH) = 2,$$

$$X(TTH) = X(THT) = X(HTT) = 1.$$

Distribution of a Random Variable

- **Definition**

The **distribution** of a random variable X on a sample space S is the set of pairs $(r, p(X = r))$ for all $r \in X(S)$, where $p(X = r)$ is the probability that X takes the value r , i.e., $p(X = r) = \sum_{s \in S, X(s)=r} p(s)$

- **Example**

Continuing with the previous example, we have

$$p(X = 3) = 1/8$$

$$p(X = 2) = 3/8$$

$$p(X = 1) = 3/8$$

$$p(X = 0) = 1/8$$

Bernoulli Trials

- **Definition:**

Each performance of an experiment with two possible outcomes is called a **Bernoulli trial**. The two possible outcomes are generally referred to as **success** and **failure**.

- **Example:**

A coin is biased so that the probability of heads is $2/3$. What is the probability that exactly four heads occur when the coin is flipped seven times?

- **Solution:**

$$C(7,4) \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^3$$

Binomial Distribution

- **Theorem**

Let X be the number of successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$. Then

$$p(X = k) = b(k: n, p) = C(n, k)p^k q^{n-k}$$

- **Definition:**

X follows the **binomial distribution** with parameters n and p , and $b(k: n, p)$ is the pmf.

- Check (from the binomial theorem)

$$\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$$

Binomial Distribution

- **Question**

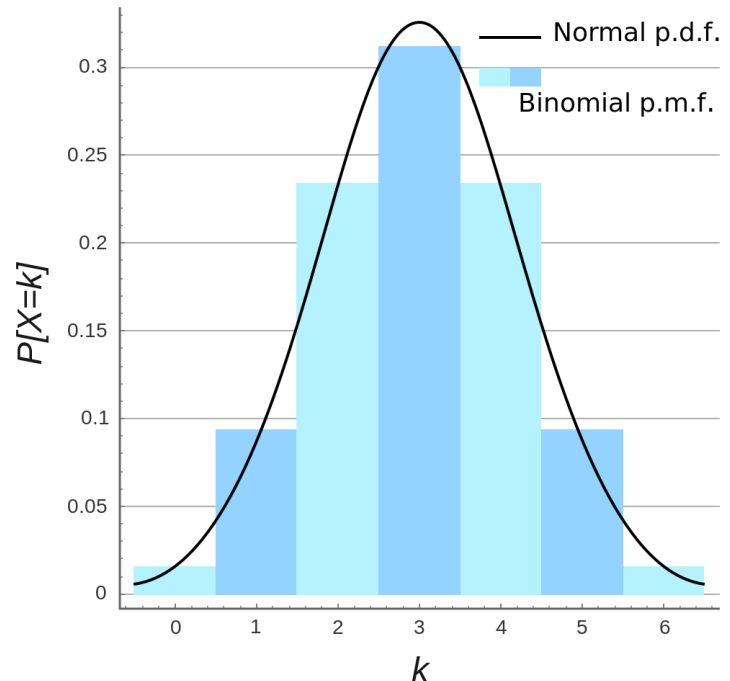
What is $\Pr(10 \leq X \leq 20)$?

- **Solution**

$$\sum_{k=10}^{20} \binom{n}{k} p^k q^{n-k},$$

which, unfortunately, has no closed form.

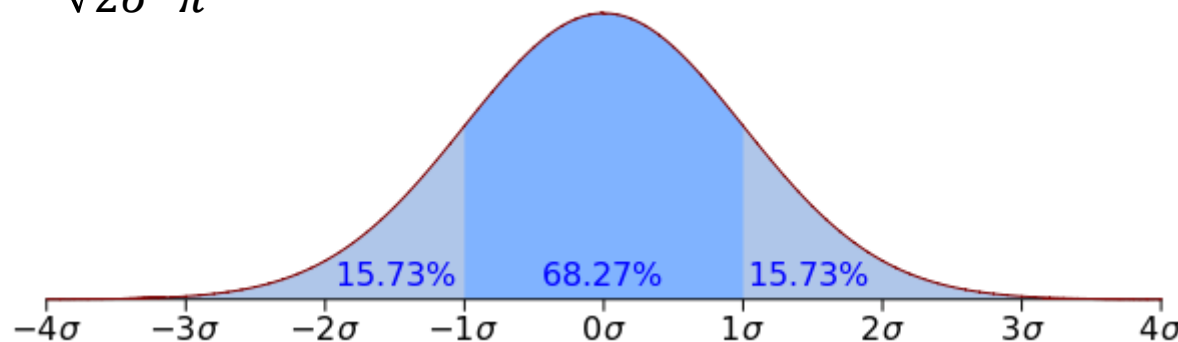
- For large n , and p not too close to 0 or 1, the normal distribution approximates binomial distribution well.
 - A special case of the central limit theorem.
 - [The Galton Board](#)



Continuous Probability Distribution

- Defined over a domain of real numbers
- Characterized by a **probability density function (pdf)**
- Example: The pdf of a normal (Gaussian) distribution

is $f(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

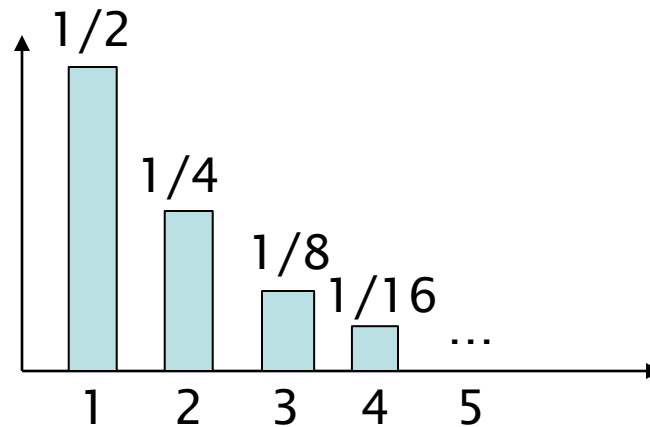


- Remark: $f(x)$ does **not** represent a probability!
- The probability of an interval: $\Pr[a \leq X \leq b] = \int_a^b f(x)dx$

Infinite Sample Space

- **Example**

Experiment: Flip a fair coin until it turns up heads.
Let random variable X = the number of flips.



- Verify $\sum_{x=1}^{\infty} p(X = x) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = 1$
- The probability of the event that we get 3 or more flips = $p(X = 3) + p(X = 4) + \dots = \frac{1}{8} + \frac{1}{16} + \dots = \frac{1}{4}$.

Geometric Distribution

- **Theorem**

Let X be the number of independent Bernoulli trials until success and the probability of success is p . Then

$$p(X = k) = (1 - p)^{k-1}p$$

- Note: X follows a **geometric distribution**
- Check: $\sum_{i=0}^{\infty} (1 - p)^i p = 1$

- **Example:**

Experiment: Roll 2 dice, probability of seeing a sum of 6 is $5/36$. Let random variable X be number of trials until getting a sum of 6.

- $p(X = 3) = \left(\frac{31}{36}\right)^2 \left(\frac{5}{36}\right)$

Expected Value

- **Definition**

The **expected value** (or **expectation** or **mean**) of a random variable X is

$$E(X) = \sum_{x \in S} p(s)X(s)$$

- **Example**

Let X be the number that comes up when a fair dice is rolled. What is the expected value of X ?

- **Solution**

The random variable X takes the values 1, 2, 3, 4, 5, or 6. Each has probability $1/6$. It follows that

$$E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \cdots + \frac{1}{6} \cdot 6 = \frac{21}{6} = \frac{7}{2}.$$

Expected Value

- Recall $p(X = r) = \sum_{s \in S, X(s)=r} p(s)$

- Theorem**

$$E(X) = \sum_{r \in X(S)} p(X = r)r$$

- Example:** Flip a coin twice. X is number of heads

s	$X(s)$	$p(s)$
HH	2	1/4
TH	1	1/4
HT	1	1/4
TT	0	1/4

r	$p(s)$
2	1/4
1	1/2
0	1/4

$$\begin{aligned}
 E(X) &= \sum_{s \in S} p(s)X(s) \\
 &= \frac{1}{4}(2) + \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(0) = 1
 \end{aligned}$$

$$\begin{aligned}
 E(X) &= \sum_{r \in X(S)} p(X = r)r \\
 &= \frac{1}{4}(2) + \frac{1}{2}(1) + \frac{1}{4}(0) = 1
 \end{aligned}$$

Expected Value

- Recall $p(X = r) = \sum_{s \in S, X(s)=r} p(s)$

- **Theorem**

$$E(X) = \sum_{r \in X(S)} p(X = r)r$$

- **Example**

Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. We know that

$$p(X = 3) = 1/8$$

$$p(X = 2) = 3/8$$

$$p(X = 1) = 3/8$$

$$p(X = 0) = 1/8$$

$$\text{So } E(X) = \frac{1}{8} \cdot 3 + \frac{3}{8} \cdot 2 + \frac{3}{8} \cdot 1 + \frac{1}{8} \cdot 0 = 1.5$$

Example: Binomial Distribution

- Recall that a random variable X with the binomial distribution with parameter n and p, q has

$$p(X = k) = C(n, k)p^k q^{n-k}$$

- We can derive its expectation through a rather complicated sequence of derivations:

$$E(X) = np$$

Linearity of Expectations

■ Theorem

If $X_i, i = 1, 2, \dots, n$ are random variables on S , and if a and b are real numbers, then

- $E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$
- $E(aX + b) = aE(X) + b$

■ Proof

$$\begin{aligned} & E(X_1 + X_2 + \dots + X_n) \\ &= \sum_{s \in S} p(s) [X_1(s) + X_2(s) + \dots + X_n(s)] \\ &= \sum_{s \in S} p(s) X_1(s) + \sum_{s \in S} p(s) X_2(s) + \dots + \sum_{s \in S} p(s) X_n(s) \\ &= E(X_1) + E(X_2) + \dots + E(X_n). \end{aligned}$$

Proof (cnt'd)

For the second result, note that

$$\begin{aligned} E(aX + b) &= \sum_{s \in S} p(s) [aX(s) + b] \\ &= a \sum_{s \in S} p(s) X(s) + b \sum_{s \in S} p(s) \\ &= a E(X) + b, \end{aligned}$$

because $\sum_{s \in S} p(s) = 1$.

Example: Binomial Distribution

- X is the number of successes in n independent Bernoulli trials, with success probability p on each trial
- $X = X_1 + X_2 + \cdots + X_n$,
where $X_i = 1$ if i th trial is successful and 0 otherwise
These X_i 's are called **indicator random variables**
- $E(X_i) = p \cdot 1 + (1 - p) \cdot 0 = p$
- $E(X) = E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = np$

Example: the Hatcheck Problem

- **Question**

A new employee started a job checking hats, but forgot to put the claim check numbers on the hats. So, the n customers just receive a random hat from those remaining. What is the expected number of hat returned correctly?

- **Solution**

Let X be the random variable that equals the number of people who receive the correct hat. We can write $X = X_1 + X_2 + \cdots + X_n$, where $X_i = 1$ if the i th person receives the correct hat and 0 otherwise. (Note that they are not independent.) $E(X_i) = 1/n$.

$$E(X) = E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = 1$$

Example: Hiring Problem

```
Hire-Assistant( $n$ ) :  
randomly permute all  $n$  candidates  
 $best \leftarrow 0$   
for  $i \leftarrow 1$  to  $n$   
    interview candidate  $i$   
    if candidate  $i$  is better than  $best$  then  
        fire  $best$   
        hire candidate  $i$   
         $best \leftarrow i$ 
```

- **Questions:** What's the expected number of hires?

Hiring Problem: Solution

- Let $X_i = 1$ if you hire candidate i and 0 otherwise.
- Let $X = \text{number of hires} = X_1 + \cdots + X_n$.
- $E(X_i) = p(X_i = 1) = 1/i$.
 - Among the first i candidates, the best has probability $1/i$ to be placed at the i th position.
- $E(X) = E(X_1) + \cdots + E(X_n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n-1} + \frac{1}{n} = \Theta(\log n)$.

Example: Balls and Bins

- **Question**

Throw m balls into n bins randomly. How many boxes will be empty in expectation?

- **Solution**

- Let $X_i = 1$ if box i is empty, and 0 otherwise.
- Let $X =$ number of empty boxes $= X_1 + \cdots + X_n$.
- $E(X_i) = p(X_i = 1) = \left(1 - \frac{1}{n}\right)^m$.
- $E(X) = E(X_1) + \cdots + E(X_n) = n \left(1 - \frac{1}{n}\right)^m$
- If throwing n balls into n bins, $n \left(1 - \frac{1}{n}\right)^n \approx \frac{n}{e}$ for large n

Average-case Analysis of Algorithms

- **Definition**

Let X be the random variable representing the running time (i.e., # of instructions) of an algorithm, when the input is drawn from a certain distribution (usually, uniform). Then the **average-case running time** of the algorithm is $E(X)$.

Example: Linear Search

```
procedure linear search(x:integer,  
                         $a_1, a_2, \dots, a_n$ : distinct integers)  
i := 1  
while (i ≤ n and x ≠ ai)  
    i := i + 1  
if i ≤ n then location := i  
else location := 0  
return location
```

- Assume that x is one of the a_i 's with equal probability
- For asymptotic analysis, it's enough to count the number of iterations, denoted as X
- $$E(X) = \frac{1}{n} \cdot 1 + \frac{1}{n} \cdot 2 + \dots + \frac{1}{n} \cdot n = \frac{n+1}{2} = \Theta(n)$$

Example: Insertion Sort

```
Insertion-Sort (A) :  
for  $j \leftarrow 2$  to  $n$  do  
     $key \leftarrow A[j]$   
     $i \leftarrow j - 1$   
    while  $i \geq 1$  and  $A[i] > key$  do  
         $A[i + 1] \leftarrow A[i]$   
         $i \leftarrow i - 1$   
    endwhile  
     $A[i + 1] \leftarrow key$   
endfor
```

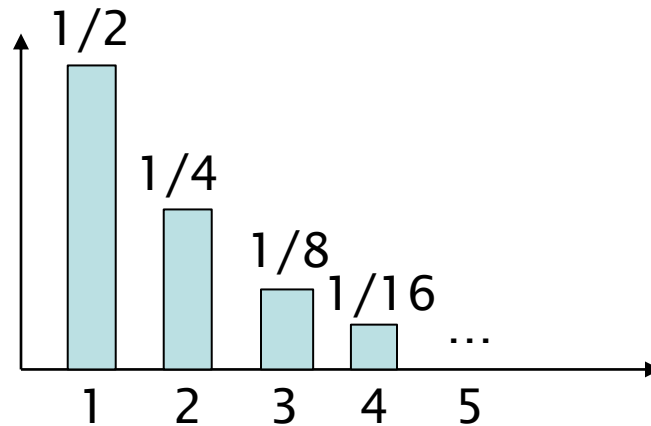
sorted	key	unsorted
--------	-----	----------

Example: Insertion Sort

- Assumption: The input is a random permutation.
- For asymptotic analysis, it's sufficient to count the total number of iterations of the inner loop, denoted X
- X_i : the # of iterations of the inner loop for inserting a_i
 - $E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)$
- a_i can be inserted into one of the first i locations with equal probability
 - The first i elements of a random permutation is a random permutation of i elements
 - The rank of the last element in a random permutation is one of $1, \dots, i$ with equal probability
- $E(X_i) = \frac{1}{i} \cdot 1 + \frac{1}{i} \cdot 2 + \dots + \frac{1}{i} \cdot i = \frac{i+1}{2}$
- $E(X) = \sum_{i=2}^n \frac{i+1}{2} = \frac{n^2+3n-4}{4} = \Theta(n^2)$

Geometric Distribution

- Experiment: Flip a fair coin until it turns up heads. Let random variable X = the number of flips.



- General case: each coin turns up heads with prob. p
- Definition**
A random variable X has **geometric distribution** with parameter $p \leq 1$ if $p(X = k) = (1 - p)^{k-1}p$, $k = 1, 2, 3, \dots$
- Theorem:** $E(X) = \sum_{k=1}^{\infty} p(1 - p)^{k-1} k = \dots = \frac{1}{p}$

Example: Coupon Collector

- **Question**

Each box of cereal contains a coupon. There are n different types of coupons. Assuming a box contains each type of coupon equally likely, how many boxes do you need to open to have at least one coupon of each type?

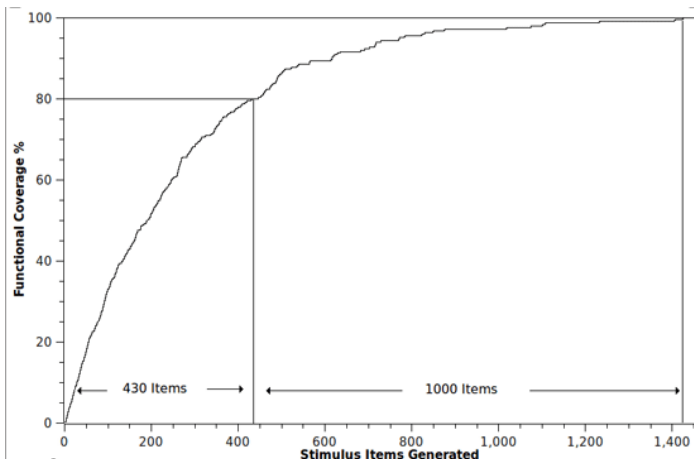


Coupon Collector: Solution

- Let X_i = number of boxes you need to open between having i and $i + 1$ distinct coupons (geometric distribution)
- Let X = number of boxes in total = $X_0 + X_1 + \dots + X_{n-1}$.

$$E[X] = \sum_{i=0}^{n-1} E[X_i] = \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{i=1}^n \frac{1}{i} = \Theta(n \log n)$$

↑
geometric distribution with $p = (n - i)/n$



Independent Random Variables

- **Definition**

The random variables X and Y on a sample space S are independent if

$$p(X = r_1 \text{ and } Y = r_2) = p(X = r_1) \cdot p(Y = r_2)$$

for all r_1, r_2

- **Theorem**

If X and Y are independent variables on a sample space S , then $E(XY) = E(X)E(Y)$.

- **Note**

Linearity of expectation holds no matter the random variables are independent or not.

Example

- Let X and Y be random variables that count the number of heads and the number of tails when a fair coin is flipped twice.
- Show that X and Y are not independent.
 - $p(X = 2 \text{ and } Y = 2) = 0$
 - $p(X = 2) = 1/4, p(Y = 2) = 1/4$
- Check
 - $E(XY) = 1 \cdot \frac{1}{2} = \frac{1}{2}$
 - $E(X) = E(Y) = 1$
 - $E(XY) \neq E(X)E(Y)$
- Check $E(X + Y) = 2 = E(X) + E(Y)$

Proof

- From the definition of expected value and because X and Y are independent random variables, it follows that

$$\begin{aligned} E(XY) &= \sum_{s \in S} X(s)Y(s) p(s) \\ &= \sum_{r_1 \in X(S), r_2 \in Y(S)} r_1 r_2 \cdot p(X(s) = r_1 \text{ and } Y(s) = r_2) \\ &= \sum_{r_1 \in X(S), r_2 \in Y(S)} r_1 r_2 \cdot p(X(s) = r_1) \cdot p(Y(s) = r_2) \\ &= \left[\sum_{r_1 \in X(S)} r_1 p(X(s) = r_1) \right] \cdot \left[\sum_{r_2 \in Y(S)} r_2 p(Y(s) = r_2) \right] \\ &= E(X)E(Y). \end{aligned}$$

Variance and Standard Deviation

- **Definition**

Let X be a random variable on the sample space S . The **variance** of X , denoted by $V(X)$ is

$$V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s) = E \left((X - E(X))^2 \right)$$

- **Definition**

The **standard deviation** of X , denoted by $\sigma(X)$, is $\sqrt{V(X)}$

- **Example**

Bernoulli trial: $X = 1$ with probability p , 0 otherwise

$$E(X) = p$$

$$V(X) = (1 - p)^2 p + (0 - p)^2 \cdot (1 - p) = (1 - p)p$$

Variance

- **Theorem**

$$V(X) = E(X^2) - E(X)^2$$

- **Proof**

$$\begin{aligned} V(X) &= \sum_{s \in S} (X(s) - E(X))^2 p(s) \\ &= \sum_{s \in S} X^2(s) p(s) - 2E(X) \sum_{s \in S} X(s) p(s) + E(X)^2 \sum_{s \in S} p(s) \\ &= E(X^2) - 2E(X)^2 + E(X)^2 \\ &= E(X^2) - E(X)^2. \end{aligned}$$

- **Corollary**

$V(aX) = a^2 V(X)$, for any real number $a \geq 0$.

Example

■ Question

What is the variance of a random variable X , where X is the number that comes up when a fair dice is rolled?

■ Solution

- $E(X) = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}$

- Using the first definition:

$$V(X) = \frac{(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2}{6} = \frac{35}{12}$$

- Using the second definition:

$$E(X^2) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6}$$

$$V(X) = E(X^2) - E(X)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Bienaymé's Formula

- **Theorem**

If X and Y are two independent random variables on a sample space S , then

$$V(X + Y) = V(X) + V(Y)$$

Furthermore, if $X_i, i = 1, 2, \dots, n$, are pairwise independent random variables on S , then

$$V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n)$$

Proof of Bienaymé's Formula

We note that

$$\begin{aligned} V(X + Y) &= E((X + Y)^2) - E(X + Y)^2 \\ &= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \end{aligned}$$

Because X and Y are independent, we have $E(XY) = E(X)E(Y)$. It follows that

$$V(X + Y) = (E(X^2) - E(X)^2) + (E(Y^2) - E(Y)^2) = V(X) + V(Y).$$

- The proof of the case with n pairwise independent random variables can be constructed by generalizing the proof of the case for two random variables.

Examples

- **Question**

X = the sum of numbers of two dice

What's $V(X)$?

- **Solution**

Write $X = X_1 + X_2$, where X_1 is the number on the first dice, X_2 is the number on the second dice

We already know $V(X_1) = V(X_2) = \frac{35}{12}$

So $V(X) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$.

- **Question**

Given that X follows the binomial distribution. What's $V(X)$?

- **Solution**

Write $X = X_1 + \cdots + X_n$, where $X_i = 1$ with prob. p

We already know $V(X_i) = p(1 - p)$

So $V(X) = np(1 - p)$.

Example

- Suppose we take a measurement by repeating the same experiment many times. Each trial returns a random measurement that is correct in expectation, but with some deviation. More precisely, let X_i be the measurement from the i -th experiments we have

$$E(X_i) = \mu, \quad V(X_i) = \sigma^2$$

If we repeat the experiments n times independently, and report the average $Y = (X_1 + \cdots + X_n)/n$. What is $E(Y)$ and $V(Y)$?

- Solution**

$$E(Y) = \frac{E(X_1) + \cdots + E(X_n)}{n} = \mu$$
$$V(Y) = \frac{V(X_1) + \cdots + V(X_n)}{n^2} = \frac{\sigma^2}{n}$$